

検索意図を考慮したナビゲーション支援システムに関する研究

著者	宋 曦媚
内容記述	筑波大学修士(情報学)学位論文 ・ 平成29年3月24日授与(37779号)
発行年	2017
URL	http://hdl.handle.net/2241/00150812

検索意図を考慮したナビゲーション支援
システムに関する研究

筑波大学

図書館情報メディア研究科

2017 年 3 月

宋曦媚

目次

図目次	iv
表目次	v
第1章 はじめに	1
1.1 研究背景	1
1.2 研究目的	3
1.3 本論文の構成	3
第2章 関連研究	4
2.1 多様化検索に関する研究	4
2.2 2画面のインタフェースに関する研究	5
2.3 意図推定に関する研究	5
第3章 サブトピック抽出と意図推定	7
3.1 中国語を対象としたサブトピックの抽出	7
3.2 意図推定	11
3.3 自動意図推定手法の評価	13
第4章 IntentNAVI システム	15
4.1 IntentNAVI システムの概要	15
4.2 サブシステム	15
4.3 2画面のインタフェースの構成	17
4.4 意図推定モジュール	20
第5章 ユーザー実験における評価	21
5.1 ユーザー実験の対象	21
5.2 ユーザー実験の課題と手続き	21
5.3 ユーザー実験の分析と結果	23
第6章 考察	33
6.1 Research Questions に関する考察	33
6.2 意図推定モジュールの失敗分析	37
6.3 IntentNAVI システムの改善に関する考察	38
第7章 おわりに	39
7.1 結論	39
7.2 今後の課題	40
参考文献	41
発表論文一覧	42
付録	43
A 最初のアンケート	43

B	検索後のアンケート	45
C	ユーザー実験における全部クエリ	48

図目次

図 1	検索を切り替えた例(Google における例)	2
図 2	インターネットの利用目的・用途(複数回答)(平成 27 年末)	2
図 3	サブトピック例のイメージ	7
図 4	サブトピック候補の取得	8
図 5	サブトピック候補から最後のサブトピックを取るまで過程	9
図 6	NTCIR12-Imine 2 タスクのクエリ例	9
図 7	予備調査で得たサブトピックの例	10
図 8	TinySVM 分類器のモデルの構築	11
図 9	意図推定用の特徴語	12
図 10	クエリ意図推定ための分類器の作業	12
図 11	5 分割交差検証の過程	14
図 12	IntentNAVI システムの全体像	15
図 13	Yahoo!Shopping の API	16
図 14	IntentNAVI システムの検索結果の画面例	17
図 15	News サブシステムの検索結果の画面例	18
図 16	Shpping サブシステムの検索結果の画面例	19
図 17	QA サブシステムの検索結果の画面例	19
図 18	検索課題の詳細	22
図 19	ユーザー実験の遂行手続き	23
図 20	意図推定群と対照群 News 課題において各サブシステムに対する平均クリック数	25
図 21	Shopping 課題において各サブシステムに対する平均クリック数	25
図 22	QA 課題において各サブシステムに対する平均クリック数	26
図 23	意図推定群のシステムが自動推定した意図	27
図 24	個別意図サブシステムに対する満足度スコアの平均値に基づいて分散分析 ..	30

表目次

表 1	NTCIR12 の Imine2 タスクにおける評価結果	10
表 2	クエリに基づいた評価結果	13
表 3	5 分割交差検証の評価結果	14
表 4	サブシステムの結果の取得	16
表 5	意図推定群と対照群 課題遂行中の各サブシステムに対する平均クリック数	24
表 6	ユーザーの切り替える件数	26
表 7	意図推定群のシステム自動判断した意図の平均回数	27
表 8	対照群のシステムランダムに選んだ意図の平均回数	28
表 9	意図推定群と対照群 アンケートによる検索意図推定の一致度スコアの平均値	28
表 10	意図推定群と対照群 サブシステムに対する満足度スコア平均値	29
表 11	意図推定群と対照群 個別意図サブシステムに対して満足した実験参加者 ..	30
表 12	全体的に最も満足したサブシステム	31
表 13	意図推定群と対照群 システム全体に対する満足度	31
表 14	個別意図に対応するサブシステムのどれかに満足する実験参加者数	33
表 15	参加者全員 全体的に最も満足したサブシステムはどれか	34
表 16	意図推定群における意図推定の正答率とシステム全体の平均満足度スコア .	35
表 17	参加者全員 個別意図サブシステムの満足度と Web サブシステムに対する満足 度	36
表 18	個別意図サブシステムごとの重み付き満足度スコア	36
表 19	Shopping 課題の失敗したクエリ (一部)	37
表 20	QA 課題の失敗したクエリ	37
表 21	失敗クエリと特徴語の頻度	38

第1章 はじめに

1.1 研究背景

ウェブの発展とともに、サーチエンジンを用いた情報検索を行う人が増えている。Sparck Jones [1]が1997年の本で「情報検索は文書検索（最近ではテキスト検索）と同義とされることが多い。すなわち、情報検索システムの仕事は、ユーザーの情報要求に適合する情報を含む文書もしくはテキストを検索することである。」と言及している。つまり、情報検索とはユーザーの情報要求に答えることがその主目的である。それはサーチエンジンにおいても同様である。

サーチエンジンを用いた情報検索を行う人が自分の情報要求をクエリとしてサーチエンジンに入力して、検索結果ではより意図に沿った検索結果を表示することが望ましい。しかし、ユーザーの検索意図は検索クエリから読み取ることが可能であるが、クエリがユーザーの意図を十分に表現できないこともある。これはウェブ検索の発展とともに、ウェブ検索における曖昧もしくは不明快なクエリが多くなったためである[2]。例えば、「オフィス」というクエリは、マイクロソフトの製品を意図しているのか、仕事を意図しているのか（曖昧なクエリ）という問題や、また、「ハリーポッター」というクエリは、本を意図しているのか、映画を意図しているのか、主人公のキャラクターを意図しているのか（不明快なクエリ）という問題が生じる[3]。このようなタイプのクエリに対する技術として、最近、多様化検索の研究が盛んである。多様化検索の目標は結果リストで多くの意図を含むようにすることである。

多様化検索では多くの意図を含んでいるものが良い結果だが、単一の意図しか持たないユーザーにとって、良い結果かどうか疑問が残る。たとえば、単一の意図を持つユーザーにとって、サーチエンジンで希望するカテゴリ(News, Shopping など)をクリックして、検索対象を切り替える場合がある。この検索対象を切り替えることを本研究ではナビゲーションと呼ぶ。Googleにおいて macbook を入力クエリとして Shopping 検索に切り替えた例は図1に示す。



図 1 検索を切り替えた例(Google における例)

本研究ではユーザー自身が検索対象を切り替える機能だけではなく、自動的にユーザー意図に応じて検索対象を切り替えることを考えて、ナビゲーションを支援する。

現実の検索対象にいろいろあり、2015 年の通信利用動向調査の結果[4]を参考にし、インターネットの利用目的・用途を探し、インターネット利用者のインターネットの利用目的・用途をみると、「電子メールの送受信」の割合が 74.3%と最も高く、「ニュースサイトの利用」(45.0%)と「商品・サービスの購入・取引」(53.3%)と「ソーシャルネットワーキングサービスの利用」(48.9%)となっている。このユーザーのインターネットの利用目的・用途データを見て、ナビゲーションの切り替え先として、本研究ではNews, Shopping, QA の 3 つの検索対象を決める。2015 年の通信利用動向調査におけるインターネットの利用目的・用途(複数回答)(平成 27 年末)を図 2 に示す。

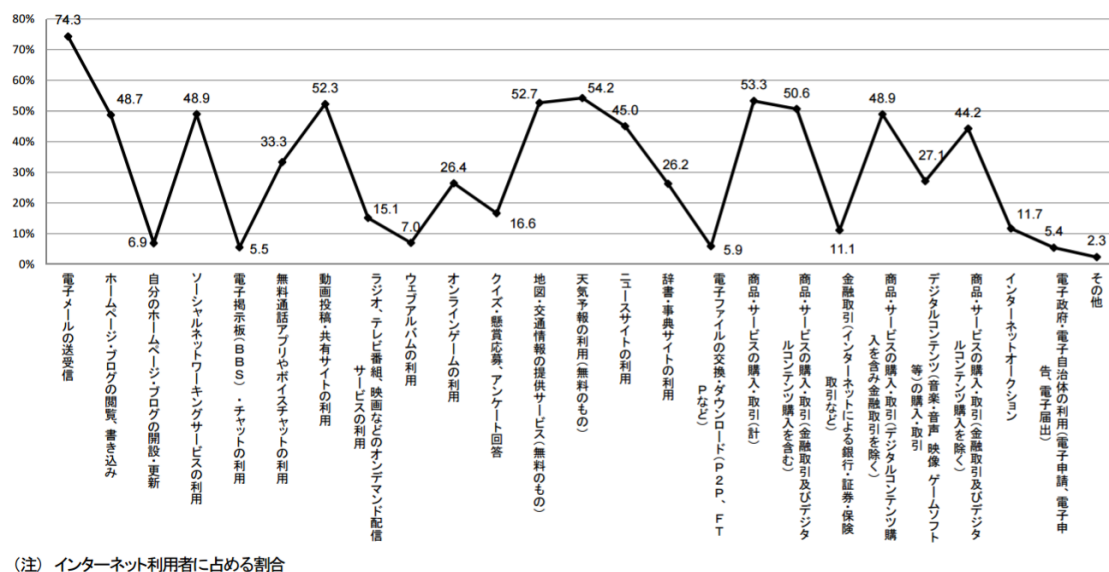


図 2 インターネットの利用目的・用途(複数回答)(平成 27 年末)

1.2 研究目的

本研究ではユーザーの意図の曖昧性にかかわらず、ユーザーを満足させることを考え、自動的にユーザー意図を推定して検索対象を切り替える機能をつけるナビゲーション支援システム(IntentNAVI)を構築する。この IntentNAVI システムを通じて、以下の Research Questions (RQs) に答える。

RQ1 : IntentNAVI システムはどれだけ有効であるか。

RQ2 : IntentNAVI システムの意図推定はどの程度正確か。

RQ3 : IntentNAVI システムの意図推定の正答率はシステム全体の有効性にどの程度影響するか。

RQ4 : IntentNAVI システムのサブシステム単体はどの程度有効か。

RQ5 : IntentNAVI システムのサブシステム単体の有効性はシステム全体の有効性にどの程度影響するか。

1.3 本論文の構成

本論文はすべて7章で構成される。第1章では、研究背景と研究目的を述べる。第2章では、関連研究のレビューを行う。第3章では、サブトピックの抽出とユーザー意図の推定について述べる。第4章では、IntentNAVI システムの構成について述べる。第5章では、IntentNAVI システムを評価するためのユーザー実験について述べる。第6章では、システムに関する考察を行う。第7章では総括を行う。

第2章 関連研究

本研究に関連する研究には多様化検索に関する研究とインタフェースに関する研究と意図推定に関する研究がある。2.1 節では多様化検索に関する研究について述べる。また、2.2 節では2画面のインタフェースに関する研究について、2.3 節では意図推定に関する研究について述べる。

2.1 多様化検索に関する研究

ウェブ検索の発展とともに、ウェブ検索における曖昧もしくは不明快なクエリが多くなったために、ユーザーの曖昧な情報要求に対する技術として、多様化検索が盛んとなってきている[2]。

いろんな研究者が検索結果の多様化に注目し、さまざまな手法が提案されている。また、評価ワークショップ TREC(Text Retrieval Conference)2012¹[5]と NTCIR-9²(NII-Test Collection for IR System)[6]においても、多様化検索タスクに着目し新しいタスクを設定して、テストコレクションの構築が始まった。アメリカの TREC と日本の NTCIR は情報検索に関連する研究分野に着目している有名なワークショップである。これらのワークショップの多様化検索に参加する研究者は多様化の検索結果を得る多数手法を提案した。

多様な検索結果を得る手法を提案した研究として Zhu ら[7]はユーザーの曖昧な情報要求に対して、relational learning-to-rank アプローチを利用した研究がある。Zhu らの手法では検索結果に対して、適合性のスコアと多様性のスコアを組み合わせでランキングすることを提案している。Plackett-Luce モデルに基づいて多様な検索結果リストを生成する。そして、TREC 評価コレクションのデータセットを利用し、評価実験を行い、効果を得た。この研究で単純に多様な結果リストを注目する。本研究では単一の意図を持つユーザーも満足するための個別意図に対応するサブシステムを含んで、個別意図に対応するサブシステムを切り替えることもできる。

従来の情報検索は検索結果の適合性(relevance)に基づいて、検索結果の有効性を評価した。なお、relevance とは、データやメッセージが探索者の情報ニーズを満たすかどうかを判定する尺度である[8]。多様化検索は検索結果の適合性だけではなく、検索結果の多様性も考える。そして、多様化検索が盛んであるため、多様化検索における評価指標も完璧になる。Sakai ら[2]は D#-measure 評価指標を提出し、他の多様化検索における評価指標(例えば、 α -nDCG and Intent-Aware など)よりメリットが多い。

多様化検索に関する研究には特殊な対象に注目し、結果の多様性を考える研究も存在する。齋藤ら[9]はEコマースにおいて複数の意図が存在する検索クエリの検索結果について、多様性を考慮したランキングを行う研究を行った。その研究で検索意図の曖昧なクエリに

¹ trec.nist.gov/pubs/trec21/t21.proceedings.html

² <http://research.nii.ac.jp/ntcir/ntcir-9/>

対して、多様性を担保しつつクエリとの関連度が低下しない手法を提案する。手法として、ユーザーがどのような意図を潜在的に持って検索したかを LDA[10]を用いて推定して、検索クエリに対してクリックした商品からそのユーザーの検索意図を表す単語行列を作成し、その単語行列を用いてトピックを推定する。推定したトピックに対して、Maximal Marginal Relevance(MMR)を利用してトピックの順序を決定する。そして先程求めた順列通りにそのトピックに属する確率の高い商品を LDA のモデルから推論し最終的な商品の順列を決定する。Yahoo!ショッピングの検索ログを用いてクエリ意図の推定と多様性の評価を行い、その有効性の検証を行った。この研究でユーザー満足するために検索結果の多様性を向上することを旨とする。本研究との違いはこの関連研究で検索結果の多様性を注目し、ユーザー意図を考慮した検索対象を切り替える機能を持ってない点がある。

2.2 2画面のインタフェースに関する研究

本研究では2画面のインタフェースを利用する。多様化検索に関するインタフェースに関する研究には Iwata らの不明確なクエリに対してユーザーができるだけ多くの適切な側面を見つけることをサポートするインタフェース(AspecTiles)[11]に関する研究が存在する。このインタフェースは2画面のインタフェースから構成され、左側に各ドキュメントのクエリの各側面に関する適合性を判断し、適合の程度を色付きタイルとして表示する。右側に Bing の API を利用した検索結果を提示した。評価実験の結果で AspecTiles は Bing より検索性能とユーザーの満足度においてよい結果を示した。各側面に関する適合性を出してユーザーをサポートする部分やインタフェースの一部を参考する。

本研究との違いはユーザー意図を推定し、検索対象を切り替える機能を持たなく、インタフェースの構成が違う。

2.3 意図推定に関する研究

ユーザーの検索意図にはさまざまな種類が存在し、検索エンジンが返す結果は必ずしもユーザーの意図を反映したものではない[12]。Trec 2012 と NTCIR9 INTENT タスクは多様化検索タスクがあり、多数の研究者を参加し、ユーザー意図推定手法を提案した。

NTCIR12-IMine2³多様化検索のタスクにおいて Nanba ら[13]が与えられたクエリに対して、サブトピックのパーティカルインテントを取得する研究が存在する。この研究では特徴語(154語)を使った分類器を用いて、サブトピックのパーティカルインテントを推定した。推定したパーティカルインテントは NTCIR12-IMine2 多様化検索タスクの評価結果で一番良い結果であった[6]。本稿では意図推定手法に Nanba らの手法を用いる。本研究との違いはこの関連研究で与えられたクエリに対してユーザー意図を考慮してサブトピックとパーティカルインテントを出す。本研究ではユーザー意図を推定してクエリに応じてサブシステムを切り替える。

TREC と NTCIR 以外の研究においても、意図推定に関する研究もある。梶並ら[12]は探索

³ <http://www.dl.kuis.kyoto-u.ac.jp/imine2/>

的情報検索を伴う多様なタスクを支援するために、ユーザーの検索意図を反映するキーワードマップと情報収集エージェントを重合したシステムアーキテクチャを提案する研究がある。この研究でユーザーの情報分析操作からユーザーの検索意図を抽出し、これに応じて情報収集・可視化を行う Poker-Maker モデルも提案する。この研究の検索意図の分析部分はユーザーの検索意図を含んだデータ分析操作(キーワードマップ上で行う)を意思決定のためにデータを多角的な視点から分析する OLAP⁴に対応づけ、3次元データキューブでキーワードマップ用データセットを表す KM キューブを提案する。この関連研究でユーザー意図を推定する機能を付けるが、対話型の探索の情報検索に注目し、提案した Poker-Maker モデルは情報収集・可視化を行う。本研究では可視化の対話型の探索に注目ではなく、ナビゲーションを支援するために検索対象を切り替える機能が重要なポイントが存在する。

⁴ <https://ja.wikipedia.org/wiki/OLAP>

第3章 サブトピック抽出と意図推定

本章ではクエリに対するサブトピックの抽出とユーザー意図の推定について述べる。サブトピックとは、クエリをより詳細したものである。例えば、「windows」というクエリに対して、「Windows 8」や「Windows update」などはサブトピックと言える。イメージは図3に示す。

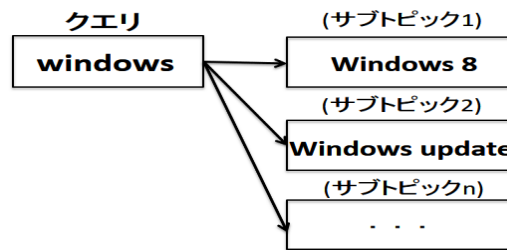


図3 サブトピック例のイメージ

3.1 中国語を対象としたサブトピックの抽出

3.1.1 概要

現実のサーチエンジンでは、ユーザーの意図を正確に把握して、ユーザーのニーズに一致するものを返すことを目指して進んでいる。そのため、ユーザー意図を明確にするために、クエリに応じたサブトピックの抽出が重要になった。私は中国語を対象としたウェブ検索の多様化について予備調査を行った。なお、本節で述べる手法と結果はNTCIR12-Imine2タスクに参加して得たものである。ただし、この手法は中国語を対象としたため、後で述べる IntentNAVI システムで利用しない。

このサブシステムの抽出は中国語を対象として NTCIR12 の Imine2 タスク（多様化検索タスク）から与えられた 100 クエリに対し、考えられる検索意図を表現した文字列（サブトピック）を生成し、ランク付けを行う。

3.1.2 サブトピックの抽出手法

サブトピックを抽出する手法として、BaiduPedia を用いて、取得した BaiduPedia の結果の中からサブトピックを取得した。手法の手続きのステップ1は与えられたクエリを BaiduPedia⁵に入力して上位10件の検索結果を出して、上位10件の検索結果のテキストを1つのテキストに纏めた。纏めたテキストの中で Stop Word を抜いて頻度が高い順から30個の名詞を取得してこの30個の名詞はサブトピック候補として取得する。なお、Stop word は「百度，百科，□吧，□片，新□，社会，特色，手机，用□，首□，分类，□□，地□，

⁵ <https://baike.baidu.com/>

音□, □片, 文□, 消息, 商城, □条, 中心, 世界, □程, 人类, 个人·」。サブトピック候補の取得を図 4 に示す。

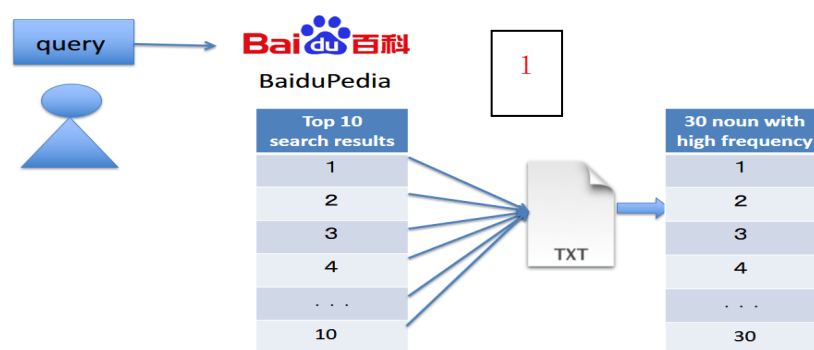


図 4 サブトピック候補の取得

ステップ 2 はサブトピック候補を抽出した後、サブトピック候補の間の類似度を計算する。サブトピックの類似度の計算手法ではサブトピック候補ごとにもう一回 BaiduPedia にクエリとしてその語を入力し、上位 10 件の検索結果を出力する。

ステップ 3 は出力結果の上位 10 件の検索結果のテキストをまとめた 1 つのテキストの中で Stop Word を抜いて頻度が高い順から 30 個の名詞とその頻度を取得し、この 30 個の名詞に基づいてサブトピック候補の間のコサイン類似度による計算する。

ステップ 4 はお互いの類似度が低いと、多様性が高くなると考えて、取得した名詞の中から類似度が低い順に 10 個の名詞を最後のサブトピックの結果として抽出する。サブトピック候補から最後のサブトピックを取るまで過程を図 5 に示す。

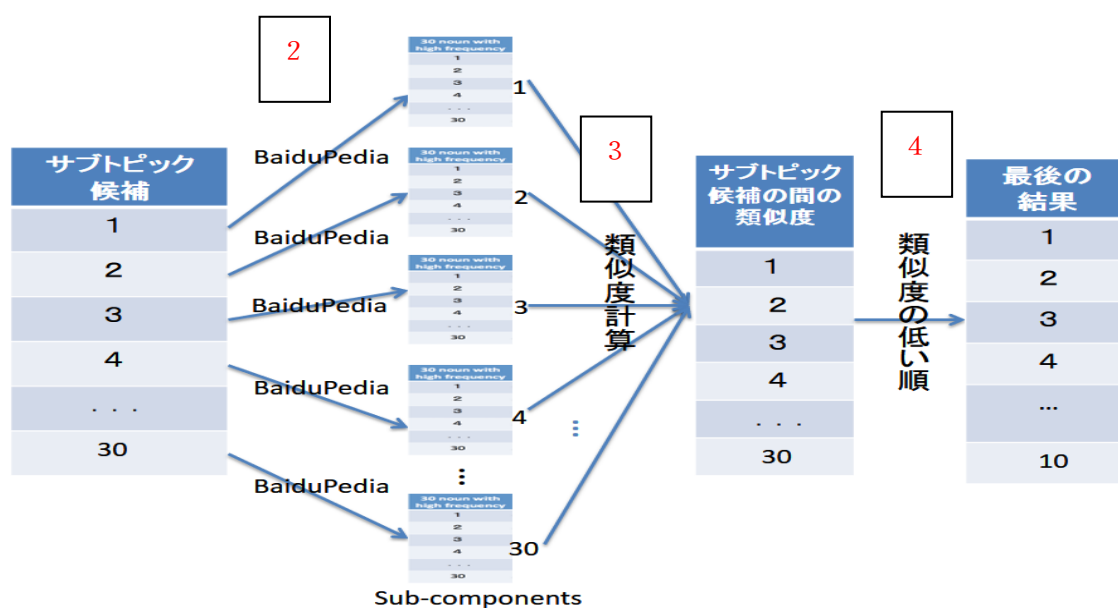


図 5 サブトピック候補から最後のサブトピックを取るまで過程

3.1.3 中国語を対象としたサブトピックの抽出の結果

NTCIR12 の Imine2 タスクのオーガナイザーからこの手法を用いて取得したサブトピックに関する評価を受け取った。NTICR-Imine2 タスクのクエリ例を図 6 に示す。

クエリ番号	中国語	日本語
IMINE2-C-001	cvcs	cvcs
IMINE2-C-002	壁紙	壁紙
IMINE2-C-003	边界	境界
IMINE2-C-004	曼彻斯特	マンチェスター
IMINE2-C-005	PS	PS
IMINE2-C-006	哀歌	悲しい歌
IMINE2-C-007	伊丽莎白	エリザベス
IMINE2-C-008	优胜美地	美しい場所
IMINE2-C-009	战斧	武器(戦争時の武器.斧)
IMINE2-C-010	星球大战	星球戦争

図 6 NTCIR12-Imine2 タスクのクエリ例

我々のグループがこの手法を用いて、取得したサブトピックの例は図 7 に示す。

	[TOPIC ID]	[SUBTOPIC]	[VERTICAL INTENT]	[SCORE]	[RUN NAME]
クエリ番号1	IMINE2-C-001	CVS美国	0.977669	IRCE-Q-C-1S	
	IMINE2-C-001	CVS药品	0.974561	IRCE-Q-C-1S	
	IMINE2-C-001	CVS数字	0.954847	IRCE-Q-C-1S	
	IMINE2-C-001	CVS代码	0.952825	IRCE-Q-C-1S	
	IMINE2-C-001	CVS公司	0.952749	IRCE-Q-C-1S	
	IMINE2-C-001	CVS版本	0.94326	IRCE-Q-C-1S	
	IMINE2-C-001	CVS文件	0.934827	IRCE-Q-C-1S	
	IMINE2-C-001	CVS服务器	0.934461	IRCE-Q-C-1S	
	IMINE2-C-001	CVS连锁	0.926255	IRCE-Q-C-1S	
	IMINE2-C-001	CVS服务	0.921835	IRCE-Q-C-1S	
クエリ番号2	IMINE2-C-002	壁纸中国	0.965423	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸墙纸	0.942151	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸宽	0.9363	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸旗航	0.933699	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸尺寸	0.9328	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸系统	0.929826	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸电脑	0.922661	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸屏	0.914942	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸图标	0.914202	IRCE-Q-C-1S	
	IMINE2-C-002	壁纸分辨率	0.913141	IRCE-Q-C-1S	

↓ ↓ ↓
 抽出した サブトピック ラン名前
 サブトピック スコア

図 7 予備調査で得たサブトピックの例

NTCIR12 の Imine2 タスクのオーガナイザーから評価結果で I-rec (再現率), D-nDCG (適合性), D#-nDCG (多様化検索における評価指標) で評価した。結果は表 1 に示す。

表 1 NTCIR12 の Imine2 タスクにおける評価結果

Topic types	I-rec@10	D-nDCG@10	D#-nDCG@10
Ambiguous	0.4827	0.4085	0.4456
Faceted	0.4650	0.4249	0.4450
Task-oriented	0.4713	0.4103	0.4408
Vertical-oriented	0.4959	0.4437	0.4698

この評価結果から見ると、どの種類のクエリにしても、D#-nDCG の評価結果の平均スコアが約 0.445 である。最高の結果を得たのは rucir グループの平均スコアが約 0.62 である。

この予備調査を通じて、ユーザー意図が多様で把握することが難しいことが分かった。予備調査で中国語を対象として、IntentNAVI システムとの言語が違い、NTCIR12 の Imine2 タスクにおける評価があまりよくなく、IntentNAVI システムでの意図推定モジュールは予備調査の手法を利用できず、日本語対象とした NTCIR12 の Imine2 タスクにおいて最高の評価結果を得た Nanba [7] らの手法を利用する。IntentNAVI システムの意図推定モジュール手

法の詳細を 3.2 節に述べる。

3.2 意図推定

ユーザーが現実のサーチエンジンにクエリを入力したとき、情報要求が曖昧なため、明確なクエリを入力することができないことがある。そして、同じトピックに対しても、ユーザーによって意図が異なる場合もある。検索システムにとって、ユーザーの検索意図を考慮する意図推定を行う必要がある。

3.2.1 意図推定手法の概要

本研究では Nanba らの意図推定手法を利用した。Nanba らの研究でユーザー意図を判断するとき、TinySVM 分類器を利用していたため、本研究でも TinySVM 分類器を利用する。TinySVM 分類器は二値分類を行うプログラムで、機械学習の一種である。二値分類とは、あるデータを入力した時、そのデータがあるジャンルに含まれそうか含まれそうではないか、ということ进行分类することである。IntentNAVI システムの個別意図に対応するサブシステムが 3 つあり、News, Shopping, QA それぞれの意図である。そのため、TinySVM 分類器で News, Shopping, QA の 3 つのモデルを用意する。分類器にデータを入力した時、入力データが News, Shopping, QA という 3 つのモデルのうちどれに属するかを判断して、入力データに一番近いモデルの名前を取り出して、ユーザー意図として推定した。TinySVM 分類器のモデルの構築は図 8 に示す。

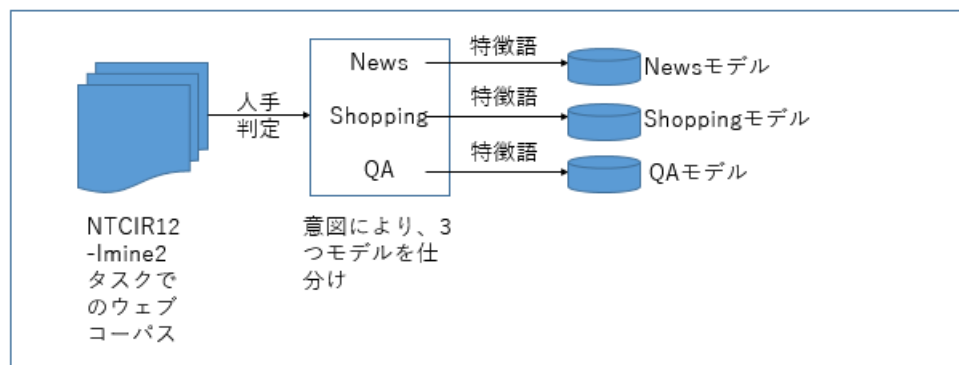


図 8 TinySVM 分類器のモデルの構築

3.2.2 クエリ意図推定のための分類器

ユーザー意図推定のための分類器を構築する。TinySVM 分類器で News, Shopping, QA といった 3 つのモデルを用意し、NTCIR12 の IMIne2 タスクでのウェブコーパスの中から抽出した 500 ファイルに基づいてトレーニングデータを作った。意図推定用の特徴語は 93 語であり、News モデルの 29 語と Shopping モデルの 41 語は Nanba らの研究の特徴語から抽出し、

Nanba らの手法で QA の特徴語が存在しなかったため、QA モデルの 23 語は自分で作成した。特徴語を図 9 に示す。

<p>News(Nanba らの研究による)</p> <p>"ニュース","新聞","トピックス","日経","速報","お知らせ","紹介","公開","配信","報じる","テレビ","記事","更新日","Introduction","Staff","Cast","Story","Character","Trailer","スタッフ","キャスト","ストーリー","あたすじ","登場人物","予告","試合結果","成績","news","headlines"</p> <p>Shopping(Nanba らの研究による)</p> <p>"販売","価格","消費税","商品","先着","限定販売","営業時間","定休日","円","受付","受け付け","日時","商品番号","数量","配送","注文","負担","カゴ","お届け","お試し","贈答カート","営業日","通販","人気","ショッピング","買取","買い取り","宅配","購読","出荷","レストラン","ランチ","グルメ","ぐるなび","本店","店舗","メニュー","amazon","rakuten.co.jp","gnavi","hotpepper"</p> <p>QA(自分で作った特徴語)</p> <p>,"質問","問題","疑問","解決","解決済み","投稿者","回答","回答数","閲覧数","悩み","教えて","知恵袋","サジェスト","chiebukuro","QA","question","answershop","qa","question","answer","chiebukuro","oshiete","q.hatena"</p>
--

図 9 意図推定用の特徴語

ウェブコーパスの中から抽出した 500 個のファイルを人力で一つずつどのモデルに属するかを判断した上で、500 ファイルで各特徴語の頻度を計算し、トレーニングデータとして利用した。現実のクエリを入力した時に、分類器の作業を図 10 に示す。

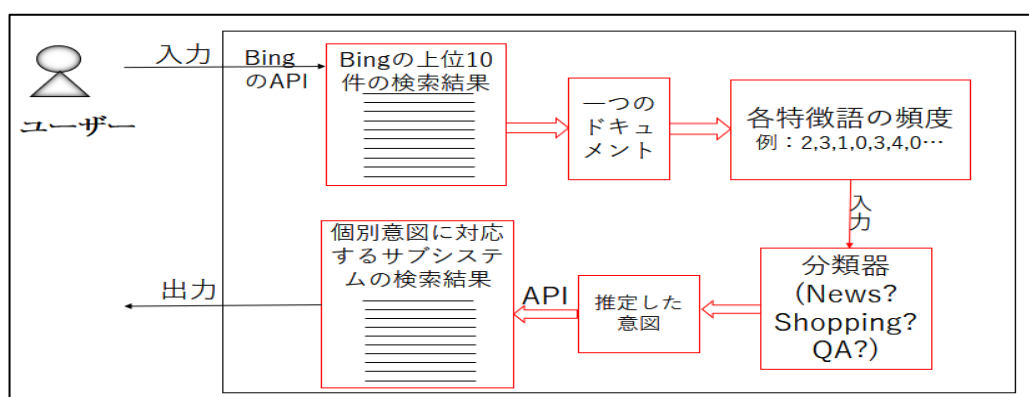


図 10 クエリ意図推定ための分類器の作業

ユーザーがクエリから意図推定にはシステムにクエリを入力した時、クエリに応じて Bing の API を用いて、Bing の上位 10 件の検索結果のテキストを 1 つドキュメントにまと

めた。まとめたドキュメントから News, Shopping, QA の各特徴語の頻度を計算して、分類器に入力し、TinySVM 分類器は二値分類器なので、News, Shopping, QA の 3 つのモデルそれぞれの出力スコア最大値となるものをクエリ意図として分類する。

3.3 自動意図推定手法の評価

本研究では Tiny 分類器を用いて、クエリに応じてユーザー意図の分類を行い、ユーザー意図を判断することにより、ユーザーに支援を行うナビゲーション支援システムを構築した。本節では構築したシステムの分類器のオフライン評価に関する結果を説明する。

3.3.1 クエリ意図推定

ここでは、NTCIR12 の IMine2 タスク (多様化検索タスク) から与えられた 100 件クエリを評価用のクエリとして用いる。NTCIR12 の IMine2 タスクで使われたクエリに関するサブトピックのパーティカルインテント毎の重要度の合計から最も大きいジャンルをそのクエリの正解ジャンルとみなして評価する。評価結果は表 2 に示す。

表 2 クエリに基づいた評価結果

クエリ	Topic (件)	分類器で分類した Topic (件)	分類器で正答した Topic (件)	精度	再現率
News	15	39	14	0.359	0.933
Shopping	40	33	24	0.727	0.600
QA	45	28	19	0.679	0.422

NTCIR12 の IMine2 タスクの評価結果で、100 件の中で 15 件のクエリの意図が News、40 件のクエリの意図が Shopping、45 件のクエリの意図が QA となっていた。TinySVM 分類器で分類した結果は、100 件のクエリの中で 39 件のクエリの意図が News、33 件のクエリの意図が Shopping、28 件のクエリの意図が QA となっていた。TinySVM 分類器で分類した結果は NTCIR12 の IMine2 タスクの評価結果と比較し、分類器の結果で正答したものは News において 14 クエリ、Shopping において 24 クエリ、QA において 19 クエリであった。

各ジャンルのクエリ数を通じ、分類器の精度と再現率を計算した。Shopping ジャンルの精度が一番高く、News ジャンルの再現率が一番高いと評価した。

3.3.2 分類器のトレーニングデータの頑健性の評価

トレーニングデータの頑健性についてモデルの評価を行い、5 分割交差検証⁶を用いて、分類器の評価を行った。評価過程は図 11 に示す。

⁶ <https://ja.wikipedia.org/wiki/%E4%BA%A4%E5%B7%AE%E6%A4%9C%E8%A8%BC>

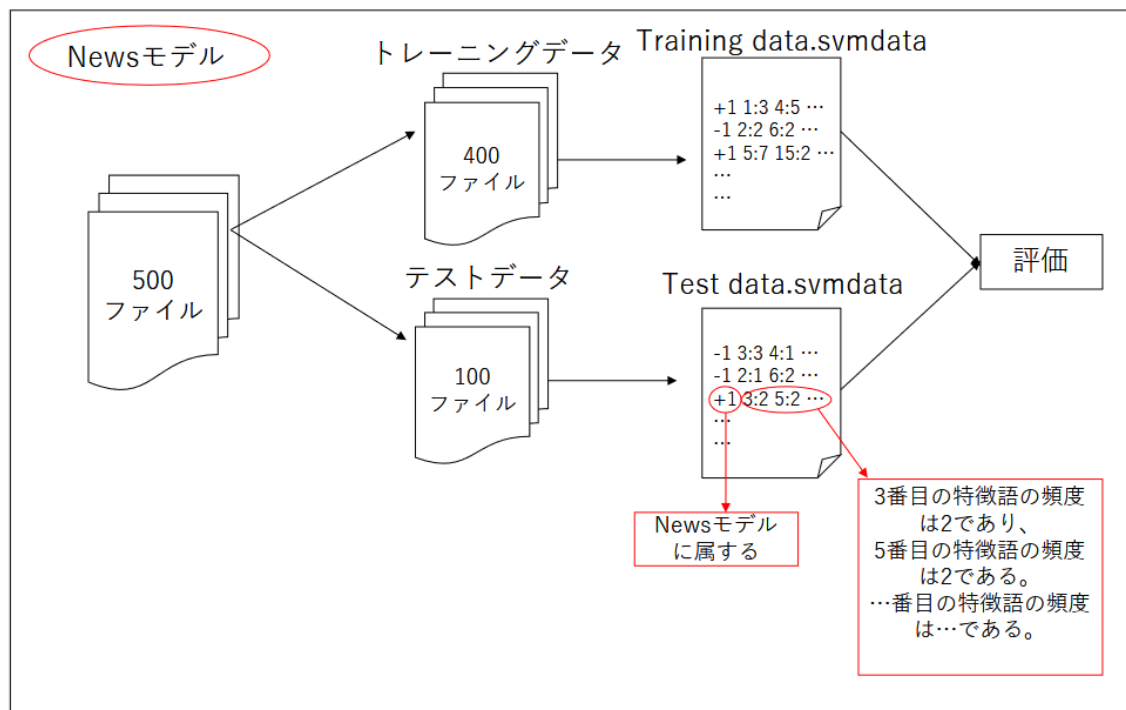


図 11 5 分割交差検証の過程

評価結果を表 3 に示す。

表 3 5 分割交差検証の評価結果

	精度	再現率	文書数
News	0.711	0.812	141
Shopping	0.919	0.842	207
QA	0.715	0.800	152
全体	0.782	0.818	500

表 3 から見ると、Shopping ジャンルのクエリは News ジャンルと QA ジャンルのクエリに比べ、精度と再現率が高くなった。クエリに基づいた評価と 5 分割交差検証評価を比べた結果、クエリに基づいた評価の方が精度と再現率が低くなった。

第4章 IntentNAVI システム

本章では IntentNAVI システムを構築する過程と手法を説明し、システムの概要について述べる。

4.1 IntentNAVI システムの概要

IntentNAVI システム全体の構成は、サブシステム、ユーザーインタフェース、意図推定モジュールの3つからなる。IntentNAVI システムの全体像は図 12 に示す。

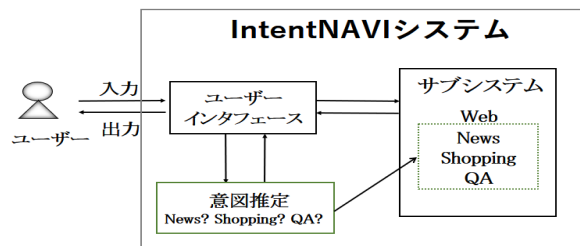


図 12 IntentNAVI システムの全体像

本研究では検索対象を切り替えるナビゲーションを支援するために、サブシステムとして News, Shopping, QA それぞれの意図に対応するサブシステムと Web サブシステムの4つを設定した。

IntentNAVI システムで2画面によるインタフェースを利用し、左側に個別意図に対応する3つのサブシステムからユーザー意図にもっとも近い1つの意図に対応するサブシステムの結果を表示し、右側に常に Web サブシステムの結果を表示する。

意図推定モジュールは IntentNAVI システムにおいてユーザー意図を考慮する部分である。意図推定モジュールではユーザーが入力したクエリに応じて、TinySVM 分類器⁷を用いて、ユーザー意図が News, Shopping, QA の中でどれかと判断し、判断された意図に対応するサブシステムの結果をインタフェースで表示する。

4.2 節で、サブシステムの詳細について述べる。4.3 節で、インタフェースの詳細について述べる。4.4 節で、意図推定モジュールについて説明する。

4.2 サブシステム

IntentNAVI システムにはサブシステムとして News サブシステム、Shopping サブシステム、QA サブシステムと Web サブシステムがある。各サブシステムが News, Shopping, QA, Web それぞれに対応するサーチエンジンの API を通じ、それぞれの API は XML 解析の結果を

⁷ <http://chasen.org/~taku/software/TinySVM/>

返すため、Nokogiri⁸を用いて XML を解析して、検索結果を得る。サブシステムの結果の取得は表 4 に示す。

表 4 サブシステムの結果の取得

サブシステム	API	取得した結果
News サブシステム	Google News	タイトル、サブタイトル、URL、写真など
Shopping サブシステム	Yahoo!Shopping	タイトル、URL、Snippet、値段など
QA サブシステム	Yahoo!知恵袋	タイトル、URL、Snippet など
Web サブシステム	Bing	タイトル、URL、Snippet など

News サブシステムでは Google News の API⁹を用いて、クエリに応じて Google News¹⁰の結果を取得する。取得した結果はタイトルやサブタイトルや URL や写真などから構成される。

Shopping サブシステムでは Yahoo!Shopping の API¹¹を用いて、クエリに応じた Yahoo!Shopping¹²の結果を取得する。取得した結果はタイトルや URL や Snippet や値段などから構成される。URL のイメージは図 13 に示す：



図 13 Yahoo!Shopping の API

QA サブシステムでは Yahoo!知恵袋の API¹³を利用し、クエリに応じて Yahoo!知恵袋¹⁴の結果を取得する。取得した結果はタイトルや URL や Snippet などから構成される。

Web サブシステムでは Bing の API¹⁵を用いて、クエリに応じて Bing¹⁶の結果を取得する。取得した結果はタイトルや URL や Snippet などから構成される。

4 つのサブシステムの取得した結果の snippet は長すぎないように、200 文字以内に確認できるよう短くした。

4 つのサブシステムではそれぞれのサーチエンジンの API を通じて、リアルタイムの上位

⁸ <http://www.nokogiri.org/>

⁹ <https://news.google.com/news?hl=ja&q=#{query}&output=rss>

¹⁰ <https://news.google.co.jp/>

¹¹ <http://shopping.yahooapis.jp/ShoppingWebService/V1/itemSearch?appid=#{}&q=#{query}>

¹² <http://shopping.yahoo.co.jp/>

¹³ <http://chiebukuro.yahooapis.jp/Chiebukuro/V1/questionSearch?appid=#{}&q=#{query}>

¹⁴ <http://chiebukuro.yahoo.co.jp/>

¹⁵ <https://www.bing.com/search?q=#{query}&format=rss&count=30>

¹⁶ <https://www.bing.com/>

30 件の結果を取得し、インタフェースに 10 件ごとの結果を 1 つのページに載せる。掲載したページのページリンクをクリックすると、ページの切り替えを行うことができる。

4.3 2 画面のインタフェースの構成

IntentNAVI システムの結果画面は 2 画面のインタフェースを利用する。IntentNAVI システムの検索結果の画面例は図 14 に示す。



図 14 IntentNAVI システムの検索結果の画面例

ユーザーがクエリを入力した際に、IntentNAVI システムがユーザー意図を推定してから、News, Shopping, QA サブシステムからユーザー意図に近い 1 つのサブシステムの結果をシステムの結果画面の左側に表示する。

ユーザーの意図は News, Shopping, QA 以外の意図を支援するため、Web サブシステムの結果を常に右側に表示する。Web サブシステムの結果が Bing の検索結果を用いるため、結果の多様化もサポートできる。

システムが推定した意図がユーザーの意図に合わない場合は、画面の上でサブシステムを News, Shopping, QA と書かれた切り替えるボタンにより、ユーザー自身がそのサブシステムを指定することで、サブシステムの切り替えを実現している。

IntentNAVI システムではサブシステムが 4 つあるが、ユーザー意図を推定するだけではなく、サブシステムの切り替える機能を合わせて用いることができる。また、意図の推定と 2 つのサブシステムの結果を取得し終えるまで待機すると、システムの処理時間が遅くなるため、プログラムでは Ajax¹⁷を使う。Ajax はウェブブラウザ内で非同期通信を行いながらインタフェースの構築を行うプログラミング手法であり、ユーザーの待ち時間 (3 秒ぐらい) を短くすることができる。本システムでは 2 つのサブシステムのうち片方の結果を得た時点でこのサブシステムの結果を表示できる。

¹⁷ <https://ja.wikipedia.org/wiki/Ajax>

熊本地震というクエリを入力した時の News サブシステムの検索結果の画面例を図 15 に示す。

IntentNAVIシステム

Search: 熊本地震 Display

news shopping QA web

news1

- <http://www3.nhk.or.jp/news/html/20170108/k10010832481000.html>
熊本地震で被災の益城町で成人式 - NHK

 **熊本地震で被災の益城町で成人式**

NHK 一連の**熊本地震**で大きな被害を受けた熊本県益城町で、8日、成人式が行われ、新成人たちが復興への決意を新たにしました。一連の**熊本地震**で大きな被害を受けた益城町では、地震から8か月余りたった今も人口の2割以上にあたるおよそ7300人が仮設住宅などでの生活 ...
[復興への決意新たに 被災地・益城町で成人式](#) [テレビ朝日](#)

[all 3 news articles >](#)

Sun, 08 Jan 2017 03:05:21 GMT

news2

- <http://www.asahi.com/articles/ASJDJ7HHRJDJTLVB016.html>
容疑者特定、決め手は片手鍋 熊本地震直後の火事場泥棒 - 朝日新聞

 **容疑者特定、決め手は片手鍋 熊本地震直後の火事場泥棒**

朝日新聞 熊本県警によると、関西にいた2人は昨年4月14日、**熊本地震**発生をラジオで聞き、翌15日、兵庫県姫路市でレンタカーを借りて熊本へ向かった。「被災地でお金を持っていそうな大きな家を狙おう」。そう話し合ったという。被災地に入ったのは本震直後の16日。高速を降りて最初 ...

Fri, 06 Jan 2017 01:24:20 GMT

news3

- <http://mainichi.jp/articles/20170107/ddl/k40/040/507000c>
「熊本の役に立つことを考え続ける」北九大が支援継続／福岡 - 毎日新聞

 **「熊本の役に立つことを考え続ける」北九大が支援継続／福岡**

毎日新聞 北九州市立大が、昨年4月の**熊本地震**で被災した自治体への支援を地道に続けている。全学組織の地域共生教育センターを中心に、発生半月後の5月上旬以降、学生延べ1000人が熊本市や益城町でボランティア活動をした他、昨年末には1〜3年生30人が以前から教員と ...

web1

- [https://ja.wikipedia.org/wiki/%E7%86%8A%E6%9C%AC%E5%9C%B0%E9%9C%87_\(2016%E5%B9%B4\)](https://ja.wikipedia.org/wiki/%E7%86%8A%E6%9C%AC%E5%9C%B0%E9%9C%87_(2016%E5%B9%B4))
熊本地震 (2016年) - Wikipedia熊本地震（くまもとじしん）は、2016年（平成28年）4月14日21時26分（日本標準時 [注釈 1]）以降に熊本県と大分県で相次いで発生している地震である。気象庁震度階級では最も大きい震度7を観測する地震が4月14日夜（前記時刻）...月, 02 1 2017 23:14:00 GMT

web2

- http://www.jma.go.jp/jp/quake/quake_local_index.html
気象庁 | 地震情報 - 気象庁 Japan Meteorological Agency地震情報の解説へよくある質問集（震度・マグニチュードなど）へ 地震資料のページへ 震度データベース検索（過去地震の検索）のページへ 緊急地震速報の発表状況へ 推計震度分布図のページへ（震度5弱以上を観測した場合に発表） 水, 04 1 2017 17:43:00 GMT

web3

- http://www.tenki.jp/bousai/earthquake/earthquake_area_by_intensity_detail?area_code=741
熊本県熊本の地震情報 - 日本気象協会 tenki.jp最新の地震情報を配信 tenki.jpの公式Twitterをご利用ください。全国の地震情報の速報をお伝えします 地震を知る 地震を知り落ち着いて行動を 地震はどうやってできるのか？ 万が一に備えて いざという時にあわてないために月, 02 1 2017 03:55:00 GMT

web4

- <http://www.asahi.com/special/kumamoto-earthquake/>
熊本地震：朝日新聞デジタル朝日新聞デジタルの熊本地震に関する特集ページです。最新ニュースのほか、支援情報などをまとめています。 ... わたしのよりどころ、大切なモノ・思い出 熊本地震半年 (10/26) 2度の震度7に襲われた熊本地震から半年が過ぎた。水, 04 1 2017 22:22:00 GMT

web5

- http://www.jma.go.jp/jma/menu/h28_kumamoto_jishin_menu.html
平成28年（2016年）熊本地震の関連情報地震関連 防災上の留意事項 平成28年（2016年）熊本地震の一連の地震活動は、全体として引き続き減衰しつつも、活動は継続しています。これまでの地震活動で揺れの強かった地域では、家屋の倒壊や土砂災害などの危険性が引き続き ...水, 04 1 2017 05:40:00 GMT

web6

- <http://news.yahoo.co.jp/story/146>
熊本県で震度6強、M7.3 その後地震相次ぐ（4月18日13時 ...16日午前1時25分ごろ熊本県で震度6強、M7.3、その後強い地震相次ぐ、死者は40人超に 【熊本一連の地震で死者計42人】熊本県を中心に続く地

図 15 News サブシステムの検索結果の画面例

冷蔵庫というクエリを入力した時の Shopping サブシステムの検索結果の画面例を図 16 に示す。

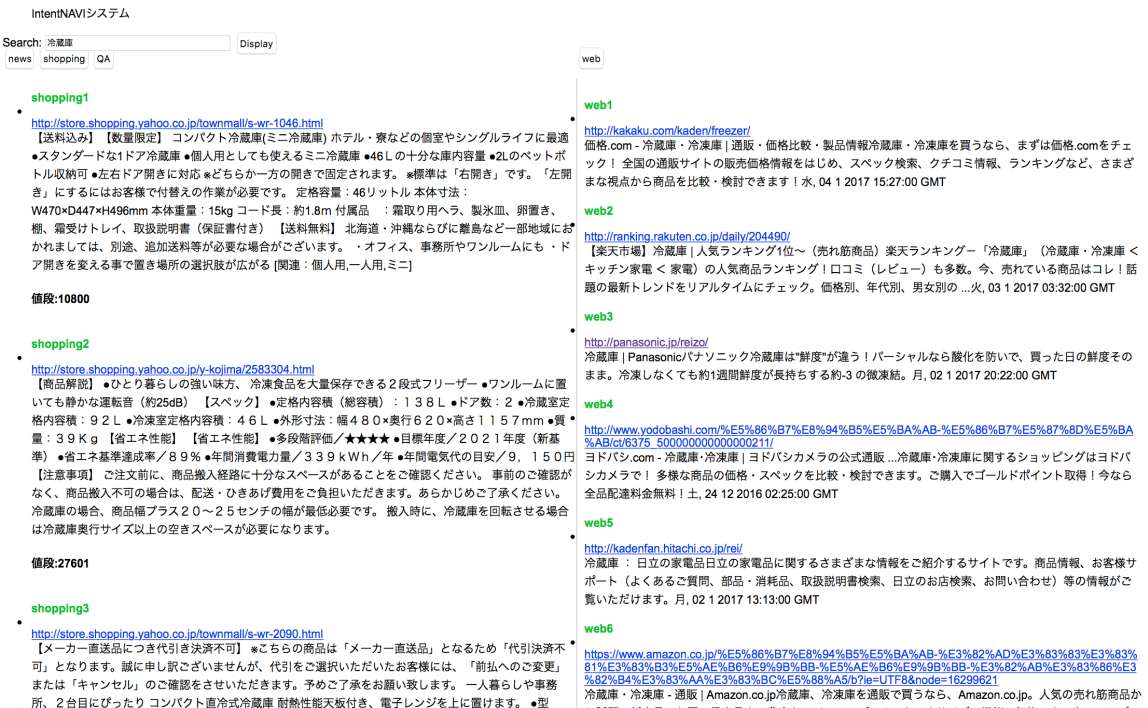


図 16 Shpping サブシステムの検索結果の画面例

エントリーシートというクエリを入力した時の QA サブシステムの検索結果の画面例を図 17 に示す。



図 17 QA サブシステムの検索結果の画面例

4.4 意図推定モジュール

ユーザー意図を考慮したナビゲーションを支援するシステムであるため、ユーザー意図を推定することがシステムにおいて重要な部分である。本研究では Nanba らの先行研究[7]でサブトピックのバーティカルインテントを取得する手法を参考にし、TinySVM 分類器を利用する。意図推定モジュールの手法の詳細に関しては 3.2 節で、TinySVM 分類器の評価の詳細に関しては 3.3 節に述べた。

第5章 ユーザー実験における評価

IntentNAVI システムの有効性を評価するため、ユーザー実験を行った。ユーザー実験の実施期間は2016年10月から12月までである。5.1節にユーザー実験の対象を述べ、5.2節にユーザー実験の課題と手続きについて述べ、5.3節にユーザー実験のデータの分析手法について述べ、5.4節にユーザー実験の結果について述べる。

5.1 ユーザー実験の対象

実験参加者は筑波大学に在籍する学生20名とした。男性が10名で女性が10名で、平均年齢は23.15歳である。参加者のうちで16名が情報学、1名が数学、2名が看護学、1名社会工学を専攻している。

5.2 ユーザー実験の課題と手続き

ユーザー実験の参加者を2グループに分けて、意図推定群の10名の参加者は通常のIntentNAVIシステムを利用し、対照群の10名の参加者は意図推定モジュールを付いていないIntentNAVIシステムを利用した。意図推定モジュールを付いていないIntentNAVIシステムはシステムの結果画面の左側の結果がNews, Shopping, QA 3つのサブシステムからランダムに1つのサブシステムの検索結果を表示する。

ユーザー実験として3つの課題を配り、News, Shopping, QA それぞれの課題はNews, Shopping, QA の3つサブシステムに対応して設定する。それぞれの課題に対する必要な情報を集めるために、IntentNAVIシステムで検索を行うよう指示した。Newsの課題は熊本地震の復興に関する情報を集めることとし、Shoppingの課題は家電製品の購入のために、家電製品に関する情報を探すこと、QAの課題は就職活動における良いエントリーシートを書くための見本などの情報を集めることとした。検索課題の詳細は図18に示す。

News 課題

2016 年 4 月 14 日に、熊本県と大分県で相次いで地震が発生しました。気象庁震度階級では最も大きい震度 7 を観測する地震が 4 月 14 日夜および 4 月 16 日未明に発生したほか、最大震度が 6 強の震度が 2 回、6 弱の震度が 3 回発生しました。一連の地震回数(M3.5 以上)は内陸型地震では 1995 年以降で最多となりました。熊本地震は被害が拡大し、復興には時間がかかりそうな状況です。被災地が日常を取り戻し、復興を遂げると、被災された地域への観光客が戻ってくることを期待できます。

あなたは、ある出版社でニュース雑誌の編集者として働いています。あなたは編集長から、このニュースを掘り下げる記事を書くための情報を集めるよう指示を受けました。編集長は、1 時間後に行われる他の部門の編集長との編集会議で記事について話し合いをする予定です。編集長は、会議でそれぞれの記事で採り上げる内容を示したいと考えています。

Shopping 課題

毎年の四月に、大学生が入学する時期になります。出身地と大学の地域が違う場合は、アパートやマンションなどを借りる必要があります。アパートで冷蔵庫やテレビなどの必要な家電製品を提供しない場合が多いです。

あなたは、北海道の出身で四月に筑波大学に入学する予定です。筑波大学の近くでアパートを借りました。アパートへの入居は二日後の予定で、このときまでに冷蔵庫などの家電製品を届けることを保証するために、早く買う必要があります。そのために、インターネットで家電製品の値段やサイズや仕様などに関する情報を多様な探す必要があります。

QA 課題

毎年、就職活動が行われています。学生が良い会社に就職するために、会社の要求に応じた書類の審査が必要です。書類選考の課題として提出を求められる採用書類をエントリーシートと呼びます。エントリーシートで自分の長所や人に負けない強みなどをうまく書けば、就職活動の役に立ちます。

あなたは、就職中の大学四年生です。ある会社の説明会に行きたいので、明日までこの会社にエントリーシートを提出しなければなりません。エントリーシートで自己 PR や自分の強み、弱みなどを書く必要があります。良いエントリーシートを書くために、内容の書き方や見本、注意点などの多様な情報を探します。

図 18 検索課題の詳細

実験参加者の主観評価を得るために、アンケートは最初のアンケートと検索後のアンケートがある。アンケートの詳細を付録[A, B]に示す。

ユーザー実験を遂行する手続きは図 19 に示す。

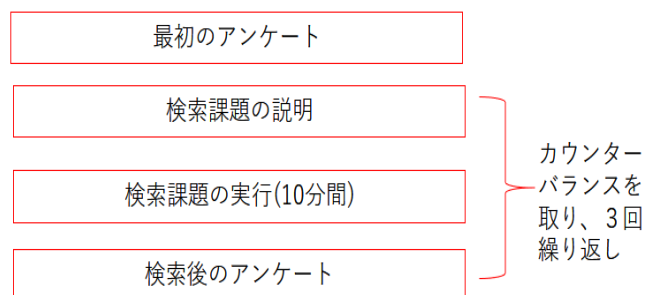


図 19 ユーザー実験の遂行手続き

ユーザー実験の初めに最初のアンケートを記入し、最初のアンケートでインターネットを使う利用状況や最近探したものなどの検索経験について質問した。最初のアンケートを記入してから、1つ目の検索課題の内容を説明し、検索を行うよう指示した。検索時間は10分で、検索の目的と合うページを見つけた際はブックマークをし、10分経過したら、1回検索を終了する。検索を行った後で、システムの有効性やユーザー満足度などを評価する検索後のアンケートを行った。各課題の遂行順序は実験参加者間でカウンターバランスを取り、3回繰り返す。

5.3 ユーザー実験の分析と結果

IntentNAVI システムの有効性などを評価するためのユーザー実験から得たデータを分析する。分析したデータはユーザー行動、意図推定の正確さと主観評価のアンケートである。データを取得する方法はアクセスログとサーバログとスクリーンキャプチャを利用する。定量的な分析では3要因分散分析を用いる。分散分析の有意水準は0.05を用いることとする。多重比較の検定にはHolm法を用いた。

5.3.1 ユーザー行動の分析

サーバログとアクセスログとスクリーンキャプチャを利用し、課題遂行中のユーザー行動を記録した。ユーザー行動は各サブシステムに対するクリック数と切り替える機能を利用した回数を含む。

意図推定群と対照群の参加者が IntentNAVI システム課題遂行中の各サブシステムに対する平均クリック数は表 5 に示す。各参加者の課題遂行中の各サブシステムに対するクリック数は記録に示す。

表 5 意図推定群と対照群 課題遂行中の各サブシステムに対する平均クリック数

課題	サブシステムクリック数(平均)									
	全体		News		Shopping		QA		Web	
	意図推定群	対照群	意図推定群	対照群	意図推定群	対照群	意図推定群	対照群	意図推定群	対照群
News	17.1	13.7	6.8	5.5	0.0	0.0	0.8	0.1	9.5	8.1
Shopping	17.6	13.0	0.4	0.8	7.4	1.7	0.2	0.1	9.6	10.4
QA	16.4	12.7	2.8	2.0	0.4	0.0	3.8	2.4	9.5	8.3

遂行中の検索課題ごとの平均クエリ数と各サブシステムに対する平均クリック数を記録した。意図推定群の遂行中の平均クリック数に基づいて、全体から見ると、どの課題にしても、Web サブシステムの結果をクリックした数は全体クリック数の半分を占めた。Web サブシステムの結果以外から見ると、News 課題と Shopping 課題においては他のサブシステムに比べて、それぞれの意図に対応するサブシステムがよく使われると分かる。対照群の遂行中の平均クリック数に基づいて、全体から見ると、どの課題にしても、Web サブシステムの結果をクリックした数は全体のクリック数の 2/3 ぐらいを占める。News 課題においては News サブシステムのクリック数は他の個別意図のサブシステムのクリック数よりやや多い。

意図推定群と対照群の各サブシステムの平均クリック数によって、分散分析を行って有意差があるかどうかを分かる。分散分析の結果から見ると、意図推定群と対照群の間に有意 ($F(1, 18)=4.97$, $p<0.05$) と見え、サブシステムの間にも有意差 ($F(3, 108)=68.54$, $p<0.01$) もあり、サブシステムと課題の間に有意差 ($F(3, 108)=11.37$, $p<0.01$) がある。意図推定群と対照群の全体のクリック数が違うためである。News 課題において分散分析の結果を図 20 に示す。

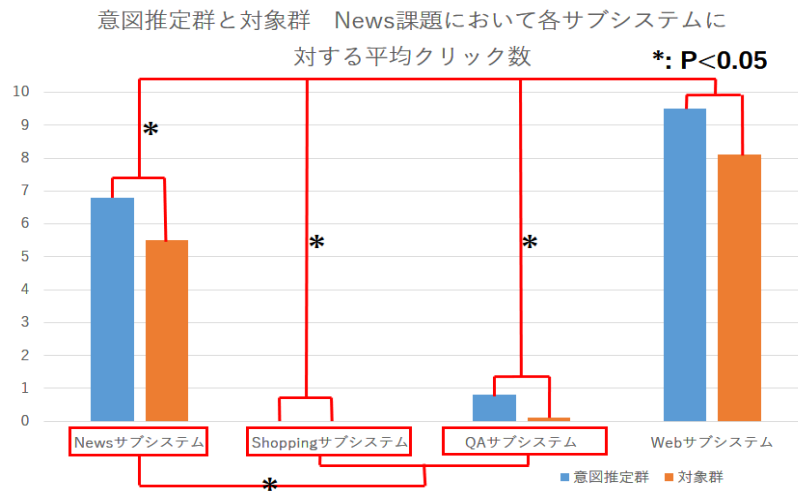


図 20 意図推定群と対照群 News 課題において各サブシステムに対する平均クリック数

News 課題において、Shopping サブシステムと QA サブシステムより、Web サブシステムと News サブシステムをよく利用すると判断されて、分散分析を行って、有意とは言える。3つの個別意図に対応するサブシステムのうち、News サブシステムが他のサブシステムより有意だとはいえる。

Shopping 課題において分散分析の結果を図 21 に示す。

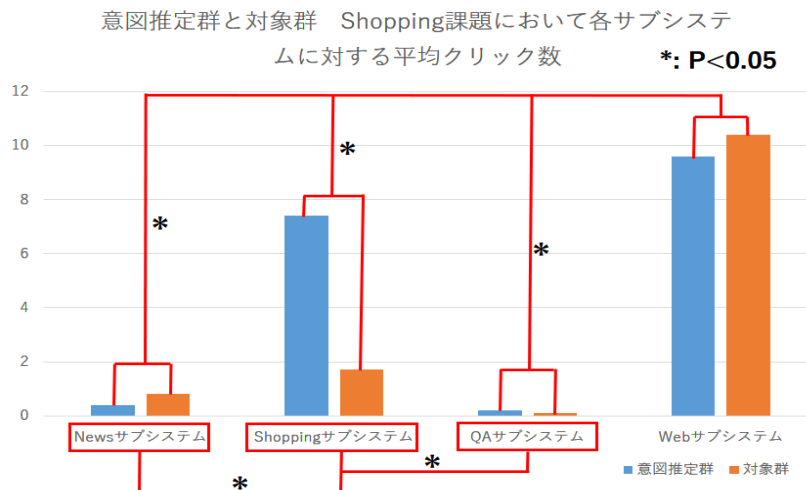


図 21 Shopping 課題において各サブシステムに対する平均クリック数

Shopping 課題において、Web サブシステムと Shopping サブシステムをよく使う。Shopping, Web サブシステムと News, QA サブシステムも有意差がある。

QA 課題において分散分析の結果を図 22 に示す。

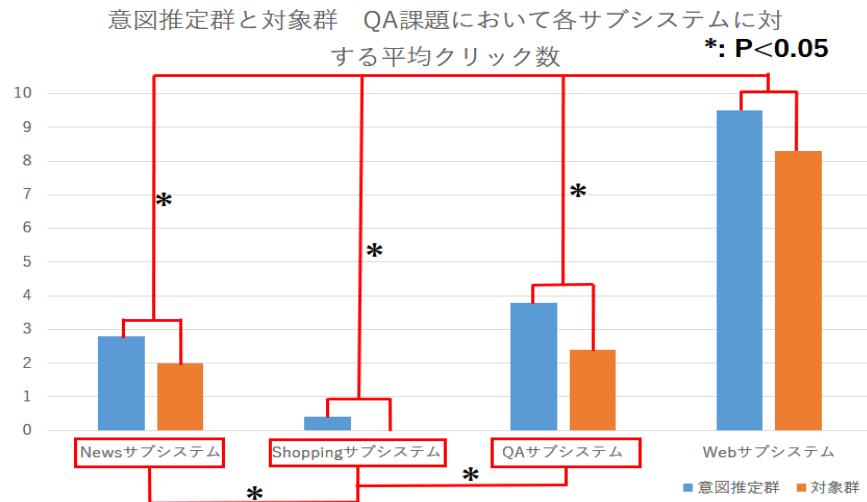


図 22 QA 課題において各サブシステムに対する平均クリック数

QA 課題において、全体から見ると、Web サブシステムをよく使う。個別意図に対応するサブシステムから見ると、Shopping サブシステムより、News サブシステムと QA サブシステムの平均クリック数が有意だといえる。

ユーザー実験におけるユーザー行動は検索結果のクリック数だけではなく、切り替えるボタンをクリックした数も重要な行動と考える。切り替える件数は表 6 に示す。

表 6 ユーザーの切り替える件数

課題	切り替える件数(平均)	
	意図推定群	対照群
News	1.7	2.8
Shopping	3.6	2.8
QA	4.6	1.7

表 6 から見ると、意図推定群において、News 課題において、切り替える件数が少なかった。News 課題において意図推定の正答率が高かったためと考える。

5.3.2 意図推定の正確さの分析

実験のログファイルを用いて、ユーザーが検索を行った際に、システムが自動的にユーザー意図を推定した回数を分析する。表 7 は意図推定群のシステム自動判断した意図の平均回数を示す。

表 7 意図推定群のシステム自動判断した意図の平均回数

課題	クエリ数 (件)	検索回数 (平均)	システム自動判断した意図(平均回数)		
			News	Shopping	QA
News 課題	4.1	8.1	7.1	0.2	0.8
Shopping 課題	6.5	12.7	2.0	8.0	2.7
QA 課題	3.5	10.9	3.4	1.1	6.4

実験中に検索を行った回数とシステムが自動推定した意図により、News 課題で平均 8.1 回検索のうちで 7.1 回 News の意図と推定し、Shopping 課題で平均 12.7 回検索のうちで 8 回 Shopping の意図と推定し、QA 課題で平均 10.9 回検索のうちで 6.4 回 QA の意図と推定した。課題に対応する意図を推定することを正答率に定義し、三つ課題の中で News 課題の正答率の方が高い。

システムが自動推定した意図の平均回数により、2 要因分散分析を行い、結果は図 23 に示す。

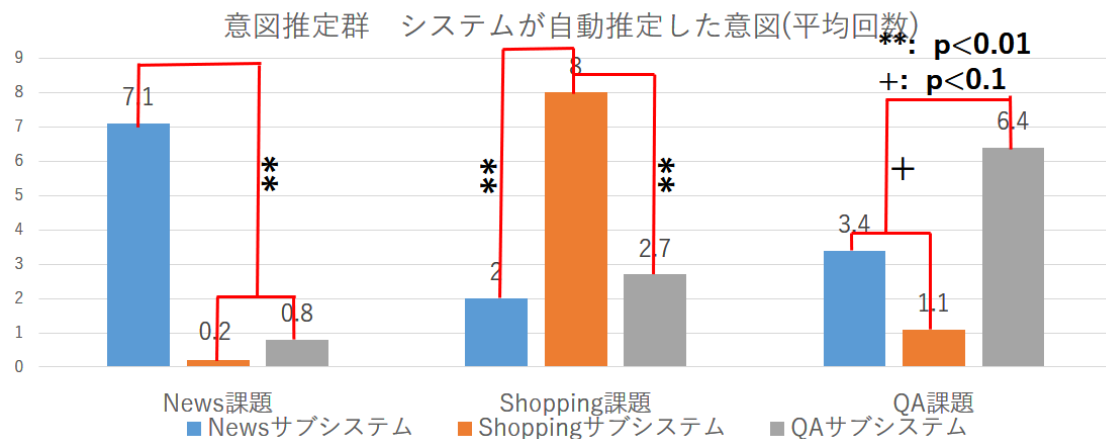


図 23 意図推定群のシステムが自動推定した意図

システムが自動推定した意図の平均回数により分散分析の結果からみると、News の課題において、News の意図推定の正確が有意 ($F(2, 18)=11.97, p<0.01$) に高いと判断した。Shopping 課題において、システムが Shopping 意図と推定された場合も多く、Shopping サブシステムと他の 2 つのサブシステムの間で有意差 ($F(2, 18)=7.60, p<0.01$) がある。QA 課題において、システムが QA 意図と推定した回数も多いであるが、分散分析を行い、有意傾向 ($F(2, 18)=3.49, p<0.1$) がある。

対照群で実験中に検索を行った回数とシステムがランダム選んだ意図の平均回数を表 8 に示す。

表 8 対照群のシステムランダムに選んだ意図の平均回数

課題	クエリ数 (件)	検索回数 (平均)	システムランダムに選んだ意図(平均回数)		
			News	Shopping	QA
News 課題	4.8	5.9	2.3	1.6	2.0
Shopping 課題	7.1	7.9	4.2	1.3	2.3
QA 課題	4.8	6.2	2.4	1.5	2.3

実験中に検索を行った回数とシステムがランダム選んだ意図により、News 課題で平均 5.9 回検索のうちで 2.3 回 News の意図と選んで、Shopping 課題で平均 7.9 回検索のうちで 1.3 回 Shopping の意図と選んで、QA 課題で平均 6.2 回検索のうちで 2.3 回 QA の意図と選んだ。Shopping 課題において、Shopping サブシステムをランダムに選んだ回数が少ない。

5.3.3 主観評価のアンケートの分析

IntentNAVI システムを利用した後、課題後のアンケートでシステムの各機能に対する満足度などを聞く。具体的なアンケート質問は付録に示す。表 9 は意図推定群と対照群の参加者のアンケートによる検索意図推定の一致度スコアの平均値である。

表 9 意図推定群と対照群 アンケートによる検索意図推定の一致度スコアの平均値

課題	意図推定群	対照群
	平均値	
News 課題	3.1	2.4
Shopping 課題	2.6	2.2
QA 課題	2.5	2.6

意図推定群のアンケートによる検索意図推定の一致度スコアの平均値から見ると、News の課題において News サブシステムの結果は一致度が高い。一方で、Shopping と QA の課題に対する推定した意図とユーザー意図の一致度がさほど高くない。

対照群のアンケートによる検索意図推定の一致度スコアの平均値から見ると、3つの課題に対して、参加者の平均一致度は「あまり一致しなかった」と「少し一致した」の間である。

意図推定群と対照群の参加者のアンケートによる検索意図推定の一致度スコアの平均値に基づいて分散分析を行い、意図推定群と対照群の間に有意差($F(1, 18)=2.56, p>0.05$)がないと言え、各課題に有意差($F(2, 36)=1.2, p>0.05$)もないと言える。

表 10 は意図推定群と対照群の参加者のサブシステム単体に対するユーザーの満足度のアンケートの結果である。

表 10 意図推定群と対照群 サブシステムに対する満足度スコア平均値

課題	意図推定群	対照群	意図推定群	対照群
	個別サブシステムの満足度		Web サブシステムの満足度	
News 課題	3.2	2.8	3.3	3.1
Shopping 課題	2.6	2.5	3.6	3.5
QA 課題	2.4	2.2	3.3	3.5

表 10 により、意図推定群は News 課題において個別サブシステムとウェブサブシステムに満足度が同じである。Shopping 課題と QA 課題は個別サブシステムの満足度より、ウェブサブシステムの満足度が高いと言える。全体から見ると、参加者は左側の結果より右側の結果の満足度が高いと言える。

対照群は、News 課題、Shopping 課題、QA 課題の 3 つ課題にしても、Web サブシステムの満足度が高いと言える。News 課題において、個別サブシステムの満足度に対する満足度がやや高いと言える。

意図推定群と対照群がサブシステムに対する満足度スコアの平均値に基づいて分散分析を行う。結果は図 24 に示す。

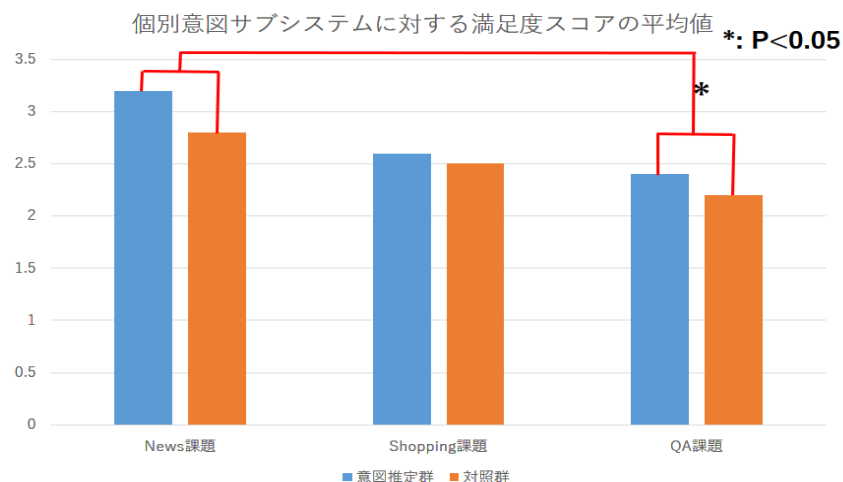


図 24 個別意図サブシステムに対する満足度スコアの平均値に基づいて分散分析

個別意図サブシステムに対する満足度スコアの平均値に基づいて分散分析を行い、意図推定群と対照群の有意差 ($F(1, 18) = 1.16, p > 0.05$) がないと言える。個別意図サブシステムの中で、QA 課題より News 課題に対する満足度スコアの方が高く、有意 ($F(2, 36) = 5.21, p < 0.05$) といえる。

Web サブシステムに対する満足度スコアの平均値に基づいて分散分析を行い、意図推定群と対照群の有意差 ($F(1, 18) = 0.02, p > 0.05$) がないと言える。各課題の間に有意差 ($F(2, 36) = 2.3, p > 0.05$) がないと言える。

もっとも満足したサブシステムがどれかを答えるために、アンケートも実施した。表 11 は意図推定群と対照群の個別意図サブシステムに対して満足した実験参加者の人数を示す。

表 11 意図推定群と対照群 個別意図サブシステムに対して満足した実験参加者

課題	News		Shopping		QA		どれも満足しない	
	意図推定群	対照群	意図推定群	対照群	意図推定群	対照群	意図推定群	対照群
News 課題	10	10	0	0	0	0	0	0
Shopping 課題	0	0	8	8	1	2	2	1
QA 課題	4	5	0	0	3	6	3	2

表 11 により、意図推定群は個別意図サブシステムに対して満足した実験参加者の人数を

整理し、個別意図のサブシステムのうち News サブシステムの結果は参加者に満足度が高いと言える。QA の課題で QA のサブシステムの結果に対する満足度があまり高くないと言える。

対照群は個別意図サブシステムに対して満足した実験参加者は News 課題において全員 News サブシステムに対する満足したと言える。Shopping 課題において Shopping サブシステムに対して満足した参加者も多い。

全体的に最も満足したサブシステムがどれかというアンケートがあり、意図推定群と対照群の参加者がサブシステム全体で最も満足したサブシステムを表 12 に示す。

表 12 全体的に最も満足したサブシステム

課題	News		Shopping		QA		Web	
	意図推定	対照群	意図推定	対照群	意図推定	対照群	意図推定	対照群
News 課題	5	8	0	0	0	0	5	2
Shopping 課題	0	0	4	1	0	0	6	9
QA 課題	2	1	0	0	1	1	7	8

News の課題において、News のサブシステムの結果に対する満足度がやや高く、意図推定群が News サブシステムの結果と Web サブシステムの結果に満足する実験参加者の人数が同じで、対照群が Web サブシステムの結果より News サブシステムの結果に満足する実験参加者の人数が多かった。意図推定群と対照群が Shopping 課題においては Shopping サブシステムより Web のサブシステムに満足する参加者が多い。QA の課題において、QA のサブシステムの結果に満足する人が少ないです。全体から見ると、ウェブサブシステムの結果が参加者の主観満足度が高いと言える。

システム全体に対する満足度を測るため、課題後のアンケートで参加者の主観評価を行う。結果は表 13 に示す。

表 13 意図推定群と対照群 システム全体に対する満足度

課題	満足度	
	意図推定群	対照群
News 課題	3.4	3.1
Shopping 課題	3.0	3.5
QA 課題	3.2	3.7

検索結果に対する満足度により、参加者全体の満足度は「少し満足している」と「とても満足している」の間である。右側のウェブサブシステムの結果にある程度満足したことにより、個別意図のサブシステムへの不満がある程度カバーされていることがわかる。また、意図推定群が News の課題のように意図推定と検索結果が良かった課題では、全体の満足度もやや向上している。

第6章 考察

IntentNAVI システムを構築した上で、ユーザー実験を行った。本章ではこれらの結果を受けて、前節までで、Research Questions (RQs) に答えるための議論を行う。6.1 節は RQ に関する考察について述べる。6.2 節は IntentNAVI システムの意図推定モジュールで失敗したクエリに関する考察について述べる。6.3 節はユーザー実験に関する考察について述べる。

6.1 Research Questions に関する考察

以下では、ユーザー実験の結果を通じて、RQs を検討する。

6.1.1 RQ1: IntentNAVI システムはどれだけ有効であるか。

IntentNAVI システムに対するユーザー実験の結果において、各サブシステムに対する満足度を主観評価として取得した。これらの満足度から IntentNAVI システムの有効性を考察する。

参加者全員が個別意図に対応するサブシステムに対して、どれかに満足する主観評価があり(複数回答可)、その結果を表 14 に示す。

表 14 個別意図に対応するサブシステムのどれかに満足する実験参加者数

課題	News サブシステム	Shopping サブシステム	QA サブシステム	どれも満足 しない
News 課題	20	0	0	0
Shopping 課題	0	16	3	3
QA 課題	9	0	9	5

個別意図に対応するサブシステムのどれかに満足する実験参加者数から見ると、News 課題において、News サブシステムの結果に対して満足する参加者が全員で、Shopping 課題において、Shopping サブシステムの結果に対する満足する参加者(16 名)が多かった。QA 課題において、News サブシステムと QA サブシステムに満足する参加者が同じ(9 名)であった。これらのデータにより、News 課題においての News サブシステムと Shopping 課題においての Shopping サブシステムに満足する参加者が多かったと言える。

全体的に最も満足したサブシステムはどれかを参加者全員が主観評価を行った結果を表 15 に示す。

表 15 参加者全員 全体的に最も満足したサブシステムはどれか

課題	News サブシステム	Shopping サブシステム	QA サブシステム	Web サブシステム
News 課題	13	0	0	7
Shopping 課題	0	5	0	15
QA 課題	3	0	2	15

全体的に最も満足したサブシステムはどれかを参加者全員が主観評価を行った。この結果により、News 課題において News サブシステムに満足した参加者が多かった。Shopping 課題と QA 課題においては Web サブシステムに満足した参加者が全員の 3/4 に占めて多かった。また、QA 課題において QA サブシステムだけではなく、News サブシステムに満足する参加者もいる。これはユーザーが QA 課題において、News サブシステムの結果も利用した可能性がある。これらにより、News 課題において News サブシステムが有効だと言えて、Shopping 課題と QA 課題においては Web サブシステムが有効だと言える。

また、上記は満足度を通じた評価だが、一方で、検索システムの有効性を図るために、検索結果のランキングと適合性に基づいて評価することも考えられる。本研究の意図推定モジュールは現実のサーチエンジン Bing の API を利用した上で、各特徴語の頻度に基づいてユーザー意図を判断した。また、IntentNAVI システムの検索結果は各サーチエンジンの API を利用し、リアルタイムの検索結果を提示する。検索結果の順番と内容が変わる可能性があるため、サブシステムの有効な評価が難しくなる。これらによる評価を行うには、評価実験時のデータ収集をより精密に行う必要があるため、今後の課題としたい。

6.1.2 RQ2: IntentNAVI システムの意図推定はどの程度正確か。

意図推定群は検索を行った回数とシステムが自動推定した意図により、News 課題で平均 8.1 回検索のうちで 7.1 回 News の意図と推定し、Shopping 課題で平均 12.7 回検索のうちで 8 回 Shopping の意図と推定し、QA 課題で平均 10.9 回検索のうちで 6.4 回 QA の意図と推定した。News 課題において正答率が 0.935 で、Shopping 課題において正答率が 0.705 で QA 課題において正答率が 0.746 である。

システムが自動推定した意図の平均回数により分散分析の結果からみると、News の課題において、News の意図と推定する回数が有意 ($F(2, 18)=11.97, p<0.01$) に多かった。Shopping 課題において、システムが Shopping 意図と推定された場合も多く、Shopping サブシステムと他の 2 つのサブシステムの間で有意差 ($F(2, 18)=7.60, p<0.01$) がある。QA 課題において、システムが QA 意図と推定した回数も多かったが、分散分析では、有意傾向 ($F(2, 18)=3.49, p<0.1$) となった。

これらの結果より、意図推定はかなり高い正答率をあげており、各課題の意図もある程度カバーできたと考えられる。意図推定の失敗分析は 6.2 節に述べる。

6.1.3 RQ3：IntentNAVI システムの意図推定の正答率はシステム全体の有効性にどの程度影響するか。

意図推定群における IntentNAVI システムの意図推定の正答率とシステム全体の満足度を表 16 に示す。

表 16 意図推定群における意図推定の正答率とシステム全体の平均満足度スコア

課題	意図推定の正答率	平均満足度スコア
News 課題	0.935	3.4
QA 課題	0.746	3.2
Shopping 課題	0.705	3.0

IntentNAVI システムの意図推定の正答率と意図推定群がシステム全体の満足度により、参加者の満足度は「少し満足している」と「とても満足している」の間であった。意図推定の正答率が一番高かったのは News 課題である。News の課題のように意図推定と検索結果が良かった課題では、全体の満足度はやや向上している。意図推定の正答率が高くなかった場合は、Web サブシステムの結果にある程度満足したことにより、全体の有効性が高かった。Web サブシステムの結果は個別意図のサブシステムへの不満がある程度カバーされていることがわかる。全体に意図推定の正答率が高いと平均満足度が高くなる傾向も示された。

6.1.4 RQ4：IntentNAVI システムのサブシステム単体はどの程度有効か。

IntentNAVI システムのサブシステム単体はどの程度有効かという RQ に対して、ユーザー実験の主観評価によって、個別意図サブシステムの満足度と Web サブシステムの満足度平均スコアと各サブシステムに対する満足度の結果から検討する。個別意図サブシステムの満足度と Web サブシステムの満足度平均スコアは表 17 に示す。

表 17 参加者全員 個別意図サブシステムの満足度と Web サブシステムに対する満足度

課題	個別意図サブシステムに対する満足度	Web サブシステムに対する満足度
News 課題	3.00	3.20
Shopping 課題	2.55	3.55
QA 課題	2.30	3.40

参加者全員が個別意図サブシステムの満足度と Web サブシステムに対する 4 段階評価に基づいた平均満足度スコアにより、News 課題において、News サブシステムがユーザーの満足度も高かった。また、Web サブシステムに対する満足度も高かった。Shopping 課題と QA 課題において個別意図サブシステムより Web サブシステムの結果の満足度スコアが高かった。全体から見ると、Web のサブシステムの結果に満足度が高いと分かる。

表 17 で述べた満足度スコアは 3 つの個別意図サブシステムの有効性をまとめているため、3 つの個別意図サブシステムそれぞれに対する満足度を重み付き満足度スコアとして表 18 に示す。これは個別意図サブシステムに対する満足度をそのサブシステムの利用頻度に応じて重み付けした値である。サブシステムごとの重み付き満足度スコアの計算方法は個別意図サブシステムのクリック数の比×参加者ごとの満足度である。

表 18 個別意図サブシステムごとの重み付き満足度スコア

課題	News サブシステム	Shopping サブシステム	QA サブシステム
News 課題	2.731	0.000	0.119
Shopping 課題	0.431	1.739	0.130
QA 課題	1.036	0.038	1.126

表 18 から見ると、News 課題において News サブシステムの個別意図サブシステムごとの重み付き満足度スコアが高かった。News 課題において、News サブシステムの有効性が高かったと言える。Shopping 課題と QA 課題において、Web サブシステムの有効性が高かったと言える。一方、QA 課題においては、正答率があまり低くないが、QA サブシステムへの重み付き満足度スコアは低かった。その原因は QA 課題において、News サブシステムと QA サブシステムと Web サブシステムの 3 つのサブシステムを利用したためである。QA 課題の曖昧性があると考え、検討する。

実験参加者のコメントから見ると、News のサブシステムで利用した結果にはタイトル、

URL, Snippet 以外にも写真やサブタイトルなどもあり、参加者にとって、News のサブシステムのような表示が見やすかった可能性がある。

検索システムの有効性を図るために、検索結果の順番と適合性に基づく方法も考えられるが、この方法は RQ1 の考察(6.1.1 節)で述べたように、今後の課題としたい。

6.1.5 RQ5 : IntentNAVI システムのサブシステム単体の有効性はシステム全体の有効性にどの程度影響するか。

サブシステム単体の有効性から見ると、News 課題において、News サブシステムの有効性が高かったと言える。Shopping 課題と QA 課題において、Web サブシステムの有効性が高かったと言える。

RQ4 の考察(6.1.4 節)に示した通り、システム全体的に満足度から見ると、参加者は個別意図のサブシステムの結果に満足する場合は個別意図のサブシステムの結果と Web のサブシステムの結果を両方利用する。個別意図のサブシステムの結果に満足しない場合は切り替える機能を利用する。一方で、個別意図のサブシステムの結果に満足しない場合、Web のサブシステムの結果を利用する場合もある。

6.2 意図推定モジュールの失敗分析

本節では、ユーザー実験において意図推定が正答しなかった失敗クエリを示し、考察する。以下では、特に 3 課題以上失敗したクエリを抜き出して分析する。News 課題の正答率が高く、失敗したクエリはあまり出なかったため、Shopping 課題と QA 課題の 3 回以上失敗したクエリを表 19、表 20 に示す。3 つの課題それぞれのクエリの詳細は付録[C]に示す。

表 19 Shopping 課題の失敗したクエリ (一部)

失敗クエリ	検索回数	推定した意図
テレビ	11	News

Shopping 課題の失敗したクエリの失敗原因を説明する。「テレビ」という失敗クエリは意図推定群の参加者が 11 回検索し、推定した意図は全部 News 意図だった。この原因はテレビが意図推定用の分類器での News モデルの特徴語であったためである。ニュースの情報源としての「テレビ」と家電製品としての「テレビ」を区別できるような手法が求められる。

表 20 QA 課題の失敗したクエリ

失敗クエリ	検索回数	推定した意図
エントリーシート書き方	7	News
エントリーシート強み 弱み	3	News
就職	3	Shopping

QA 課題の失敗クエリ(表 20)に対する意図推定モジュールでの各特徴語の頻度を表 21 に示す。システムの実施時点によって、各特徴語の頻度は変わる可能性がある。表 21 に示したデータは 2017 年 1 月に作ったものである。

表 21 失敗クエリと特徴語の頻度

失敗クエリ	特徴語と特徴語の頻度
エントリーシート書き方	news : 3283、記事 : 2421
エントリーシート強み 弱み	news : 2160、記事 : 1880、問題 : 1395
就職	news : 1823、記事 : 893、円 : 781

この失敗クエリと対応する特徴語の頻度を見ると、特徴語の頻度の高いものに、News や Shopping の特徴語があることがわかる。これにより、推定した意図が間違えたと分かる。

6.3 IntentNAVI システムの改善に関する考察

ユーザー実験の課題後のアンケートにおいて自由記述コメントがいくつか得られたため、コメントに対して、IntentNAVI システムの改善点を検討する。

Shopping の課題において、買い物を探す場合は、ユーザーにとって、買い物の値段の安い順などの順番を付ける結果リストが欲しいという参加者からのコメントがあった。また、商品のイメージを返したら、便利だというコメントもあった。本研究では Yahoo!Shopping の API を直接利用し、結果を取り出しているが、今後の課題として、Shopping のサブシステムの結果に商品の写真などを付けたり、安い順や更新順などにより、結果リストをリランクすることが考えられる。

IntentNAVI の検索結果においてタイトル、URL、Snippet などがあり、検索結果で文字ばかりだと感じる参加者もいた。IntentNAVI の検索結果にはタイトル、URL, Snippet 以外にも写真やサブタイトルなどを付けると、参加者にとって、検索結果が見やすくなる可能性がある。

第7章 おわりに

本研究ではユーザーの意図の曖昧性にかかわらず、ユーザーを満足させるためのナビゲーション支援システムを構築した。ナビゲーション支援システムを開発手法と IntentNAVI システムの有効性に評価するためのユーザー実験を実施した。

7.1 結論

ユーザー実験を通じて、IntentNAVI システムの意図推定の手法を利用して、News 課題において正答率が 0.935 で、Shopping 課題において正答率が 0.705 で QA 課題において正答率が 0.746 である。全体の正答率はある程度の精度だったが、News 課題が一番高かった。4 つのサブシステムの全体から見ると、ウェブのサブシステムの結果が参加者の主観満足度が高いと言え、左側の 3 つのサブシステムのうち、News 課題においては News のサブシステムに対して、ユーザーの満足度が高い。一方で、QA のサブシステムの満足度とクエリ意図との一致度が高くない。システム全体として、参加者は個別意図のサブシステムの結果に満足する場合は個別意図のサブシステムの結果と Web のサブシステムの結果を両方利用する。個別意図のサブシステムの結果に満足しない場合は切り替える機能を利用する行動が見られた。一方で、個別意図のサブシステムの結果に満足しない場合、Web のサブシステムの結果を利用するようになる傾向もあった。

ユーザー実験から得たデータから見ると、意図推定の正答率が高いと、全体の満足度が高くなる傾向があり、意図推定モジュールの正確さが全体の満足度に影響があると言える。本研究では意図推定モジュールで News 課題において、意図推定の正答率が高く、ナビゲーションをよく支援できる。一方、QA 課題においては、正答率があまり低くないが、QA サブシステムへの満足度は低かった。その原因は QA 課題において、News サブシステムと QA サブサブシステムと Web サブシステムの 3 つのサブシステムを利用したためである。

個別意図に対応するサブシステムが 3 つあり、意図推定の正答率が高い場合にはそのサブシステムにより満足される一方で、意図推定の正答率が低い場合には、切り替えボタンを利用し、満足度もほとんど下がらなかった。意図推定と切り替えボタンによる手法により、ナビゲーションを支援できると言える。

本研究で利用した 2 画面のインタフェースについて、実験参加者の検索行動(結果のクリック数)から見ると、個別意図に対応するサブシステムと Web サブシステム両方も利用した。実験参加者にとって、2 画面のインタフェースもある程度に働いたといえる。

News 課題において News サブシステムを利用した結果に対する満足度が高かったのは Google News の API を用いて取得した結果にタイトル、URL、Snippet 以外にも写真やサブタイトルなどもあり、参加者にとって、このような表示が見やすかった可能性がある。

7.2 今後の課題

IntentNAVI システムはユーザーの意図に応じて検索対象を切り替えることを支援するためのシステムである。今後の課題には、(1)もっと多くの種類のサブシステムを入れること、(2)各サブシステムから得られた結果リストをユーザーのニーズに合う順番で表示すること、(3)見やすいインタフェースに取り組むこと、(4)有効性評価を行うことがある。

本研究では検索対象は News, Shopping, QA が三つあり、現実のユーザーの意図に対応するにはサブシステムの数が不十分だと考えられる。例えば、Google には画像や動画などの切り替えボタンを持つ。また、NTCIR-12 IMine2 タスクでも、百科事典、画像などの意図を対象としているため、これらは追加すべき意図の候補となりうる。このような新しい意図を追加するには、意図推定モジュールにおける分類器モデルの追加と検索サブシステムの追加が必要となる。

また、Shopping の課題において、買い物を探す場合は、ユーザーにとって、買い物の値段の安い順などの順番を付ける結果リストが欲しいという参加者からのコメントがあった。商品のイメージを返したら、便利だというコメントもあった。本研究では Yahoo!Shopping の API を直接利用し、結果を取り出しているが、今後の課題として、Shopping のサブシステムの結果に商品の写真などを付け、安い順や更新順などにより、結果リストをリランクすることが考えられる。このような個別意図サブシステムの対象に合わせた並べ替えも今後の課題となる。News サブシステムで見られたように、サムネイル画像などの表示は探索を支援すると考えられるため、各個別意図ごとに表示内容を工夫することも今後の課題である。

本研究の意図推定モジュールは現実のサーチエンジン Bing の API を利用した上で、各特徴語の頻度に基づいてユーザー意図を判断した。また、IntentNAVI システムの検索結果は各サーチエンジンの API を利用し、リアルタイムの検索結果を提示する。検索結果の順番と内容が変わる可能性があるため、サブシステムの有効な評価が難しくなる。現時点はユーザーの満足度を有効な指標として評価する。リアルタイムの結果を用いた有効性評価は今後の課題としたい。

本研究ではユーザー実験の検索課題が News, Shopping, QA それぞれの意図に対応して設定した。情報収集なタスクのため、ユーザーが入力したクエリは自由である。この入力したクエリは意図を推定して個別意図に合致するかを疑問する。検索課題と意図の関係の繋がりを今後の課題もしたい。

参考文献

- [1] Karen Sparck Jones and Peter Willet. Readings in Information Retrieval. 1997.
- [2] 情報アクセス評価方法論：検索エンジンの進歩のために. 酒井哲也著. コロナ社, 2015. 6.
- [3] Rakesh Agrawal, Sreenivas Gollapudi, Alan Halverson, Samuel Ieong. Diversifying search results. Proceeding WSDM' 09. pp.5-14. 2009.
- [4] 平成 27 年通信利用動向調査の結果. 総務省. 2016 年 7 月 22 日.
http://www.soumu.go.jp/johotsusintokei/statistics/data/160722_1.pdf
- [5] Charles L.A. Clarke, Nick Craswell, Ellen M. Voorhees. Overview of the TREC 2012 Web Track. Proceeding TREC , pp.1-8. 2012.
- [6] Takehiro Yamamoto, Yiqun liu, Min Zhang, Zhicheng Dou, Ke Zhou, Ilya Markov, Makoto P.Kato, Hiroaki Ohshima and Sumio Fujita. Overview of the NTCIR-12 IMine-2 Task. Proceeding 12th NTCIR, pp.8-26. 2016.
- [7] Yadong Zhu, Yanyan Lan, Jiafeng Guo, Xueqi Cheng, and Shuzi Niu. Learning for search result diversification. In Proceedings of the 37th International ACM SIGIR Conference on Research & development in information retrieval. pp. 293-302, 2014.
- [8] 日本図書館情報学会研究委員会編. 情報アクセスの新たな展開. 勉誠出版. pp.52. 2009.
- [9] 齋藤 祐樹, 田頭 幸浩, 小野 真吾, 塚本 浩司. 意図推定に基づく検索結果の多様性の向上. DEIM Forum 2015 A1-6.
- [10] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. Latent dirichlet allocation. J. Mach. Learn. Res., Vol.3, pp.993-1022, March 2003.
- [11] Mayu Iwata et al. AspecTiles : Tile-based Visualization of Diversified Web Search Results. Proceedings of SIGIR' 12, pp. 85-94. 2012.
- [12] 梶並知記, 高間康史. Poker-Maker モデル：ユーザの検索意図を反映するキーワードマップと情報収集エージェントの連携による探索的情報検索. 情報知識学会誌, Vol.20, No.3, pp.277-292, 2010.
- [13] Hidetsugu Nanba, Tetsuya Sakai, Noriko Kando, Atsushi Keyaki, Koji Eguchi, Kenji Hatano, Toshiyuki Shimizu, Yu Hirate, and Atsushi Fujii. NEXTI at NTCIR-12 Imine-2 Task. IN Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies, pp. 27-30, 2016.

発表論文一覧

Ximei Song, Yuka Egusa, Hitomi Saito and Masao Takaku. IRCE at the NTCIR-12 Task. Proceedings of the 12th NTCIR Conference. pp. 55-59. 2016.

宋曦媚, 高久雅生. 検索意図を考慮したナビゲーション支援システム. 第9回 Web インテリジェンスとインタラクション研究会, 東京, pp. 1-6. 2016.

付録

A 最初のアンケート

最初のアンケート

被験者 ID :

氏名 :

以下の問の該当する答えに○を付けてください。

問 1 : あなたの年齢は_____。

問 2 : あなたの性別を選んでください。

男

女

問 3 : サーチエンジンを初めて利用したのはいつ頃ですか。該当する年を選んでください。

1995 年以前	1996	1997	1998	1999	2000			
2001	2002	2003	2004	2005	2006	2007	2008	2009
2010								
2011	2012	2013	2014	わからない				

問 4 : 最近サーチエンジンで探したことを 5 つ挙げてください。

問 5 : 過去 1 年間に利用した情報通信機器はどれですか。該当するものすべてに丸をつけてください。

- ☐ 携帯電話（スマートフォンを含む）
- ☐ 携帯情報端末（iPad など）
- ☐ パソコン
- ☐ インターネットに接続できるテレビ
- ☐ インターネット対応型家庭用ゲーム機

- 問6：インターネットを利用した方全員にお尋ねします。過去1年間にどのような目的でインターネットを利用しましたか。該当するものすべてをチェックしてください。

- 44

B 検索後のアンケート

検索後のアンケート

被験者 ID :

氏名 :

課題番号 : 1 2 3

以下の問の該当する答えを選択してください。

問 1 : あなたは今回探そうとする課題についての知識をどれくらい持っていましたか。

- ☐ まったく持っていなかった
- ☐ 少し持っていた
- ☐ 持っていた
- ☐ 十分に持っていた

問 2 : あなたにとって、この課題の内容は興味のあるものでしたか。

- ☐ 興味がない
- ☐ やや興味がない
- ☐ 少し興味がある
- ☐ とても興味がある

問 3 : あなたにとって、この課題について情報を収集するのは簡単でしたか。

- ☐ とても簡単だった
- ☐ 少し簡単だった
- ☐ 少し難しかった
- ☐ とても難しかった

問 4 : あなたは課題作成に必要な情報を十分に収集できたと思いますか。

- ☐ まったく思わない
- ☐ あまり思わない
- ☐ 少しそう思う
- ☐ 非常にそう思う

問 5：必要な情報を収集するために十分な時間はありましたか。

- ☐ まったくなかった
- ☐ あまりなかった
- ☐ 少しあった
- ☐ 十分あった

問 6：検索時に左側のサブシステムに示した情報はあなたの検索意図と一致しましたか。

- ☐ まったく一致しなかった
- ☐ あまり一致しなかった
- ☐ 少し一致した
- ☐ 完全に一致した

問 7：左側のサブシステムを切り替える前で左側に示した情報に満足しましたか。

- ☐ 非常に不満である
- ☐ 少し不満である
- ☐ 少し満足している
- ☐ とても満足している

問 8：左側のサブシステムでの切り替えるボタンを利用しましたか。

- ☐ 利用した
- ☐ 利用しなかった

問 9：左側のサブシステムの中で、あなたはどのサブシステム (News, Shopping, QA) から示した情報に満足しましたか。(複数選択可能)

- ☐ News
- ☐ Shopping
- ☐ QA
- ☐どれも満足しなかった

問 10：左側のサブシステムを切り替える機能は情報を収集する役に立ちましたか。

- ☐ まったく役に立たなかった
- ☐ あまり役に立たなかった
- ☐ 少し役にたった
- ☐ とても役に立った

問 11：右側のサブシステムから示した情報に満足しましたか。

- ☐ 非常に不満である
- ☐ 少し不満である
- ☐ 少し満足している
- ☐ とても満足している

問 12：ナビゲーション支援システムの全体で、あなたはどのサブシステム (News, Shopping, QA, Web) から示した情報にもっとも満足しましたか。

- ☐ News
- ☐ Shopping
- ☐ QA
- ☐ Web

問 13：示された検索結果は全体として見て、満足できるものでしたか。

- ☐ 非常に不満である
- ☐ 少し不満である
- ☐ 少し満足している
- ☐ とても満足している

問 14：検索中に気付いた点、システムの改善点など、もしあれば、自由に記入してください。

C ユーザー実験における全部クエリ

意図推定群のクエリ：

参加者 ID	課題	クエリ	推定した意図
1	News	熊本県	News
1	News	熊本地震の被害様子	News
1	News	熊本地震復興	News
1	News	熊本現状	News
2	News	熊本 復興	News
2	News	熊本 復興 作業	News
2	News	熊本 被害様子	News
2	News	熊本 被害	News
3	News	熊本地震 復興	News
3	News	熊本城 地震 修復	News
3	News	熊本地震	News
3	News	熊本地震 被害まとめ	News
4	News	熊本地震 被害	News
4	News	熊本地震 復興	News
4	News	熊本地震 被害 写真	News
4	News	熊本地震 復興 写真	News
5	News	熊本地震 復興状況	News
5	News	熊本地震 復興	News
5	News	熊本地震 現状	News
5	News	熊本地震 災害状況	News
5	News	熊本地震 被害状況	News
5	News	熊本地震 被害状況 熊本城	News
5	News	熊本 観光地 被害状況	News
5	News	阿蘇神社 被害状況	News
5	News	阿蘇神社 被害	News
5	News	熊本 犠牲者	News
5	News	熊本 地震 リアルタイム」	News
5	News	熊本 地震 リアルタイム	News
5	News	熊本 地震	News
6	News	熊本	News
6	News	熊本 被害状況	QA
6	News	熊本 復興	News

6	News	熊本 今	News
7	News	熊本地震	News
7	News	熊本地震 状況	News
7	News	熊本地震 被害状況	News
8	News	地震 復興	News
8	News	地震 復興 熊本	News
9	News	熊本地震	News
9	News	熊本地震 現在	Shopping
9	News	熊本地震の被害様子	News
10	News	熊本 地震 復興	News
10	News	熊本 地震 現状	News
10	News	熊本 現状 様子	QA
10	News	熊本 地震 現状 様子	News
1	Shopping	テレビ	News
1	Shopping	冷蔵庫	Shopping
1	Shopping	レンジ	Shopping
2	Shopping	家電製品	Shopping
2	Shopping	冷蔵庫 筑波	Shopping
2	Shopping	家電製品 筑波	Shopping
2	Shopping	テレビ	News
3	Shopping	つくば 家電	Shopping
3	Shopping	一人暮らし 家電 サイズ	QA
3	Shopping	一人暮らし 冷蔵庫 大きさ	QA
3	Shopping	一人暮らし 電子レンジ	Shopping
3	Shopping	一人暮らし パソコン 大学生活	News
3	Shopping	一人暮らし ノートパソコン 価格帯	QA
3	Shopping	一人暮らし ポット	QA
3	Shopping	一人暮らし ih	QA
4	Shopping	テレビ 一人暮らし おすすめ 大きさ	QA
4	Shopping	家電 一人暮らし おすすめ 大きさ	QA
4	Shopping	炊飯器 一人暮らし おすすめ	Shopping
4	Shopping	冷蔵庫	Shopping
4	Shopping	洗濯機	Shopping
5	Shopping	dell パソコン	QA
5	Shopping	冷蔵庫 価格.com	Shopping

5	Shopping	冷蔵庫 amazon	Shopping
5	Shopping	電子レンジ amazon	Shopping
5	Shopping	洗濯機 amazon	Shopping
5	Shopping	テレビ amazon	News
5	Shopping	レコーダー amazon	Shopping
5	Shopping	エアコン amazon	Shopping
5	Shopping	座椅子 amazon	Shopping
5	Shopping	デスク amazon	Shopping
5	Shopping	デスク ニトリ	Shopping
5	Shopping	パソコンチェア ニトリ	Shopping
5	Shopping	本棚 ニトリ	Shopping
5	Shopping	カラーボックス ニトリ	Shopping
5	Shopping	テーブル ニトリ	Shopping
5	Shopping	丸テーブル ニトリ	Shopping
5	Shopping	折りたたみテーブル ニトリ	Shopping
5	Shopping	カーテン ニトリ	Shopping
6	Shopping	大型テレビ	Shopping
6	Shopping	パソコン	Shopping
6	Shopping	dell パソコン	Shopping
6	Shopping	dell パソコン	QA
7	Shopping	家電	Shopping
7	Shopping	家電 冷蔵庫	Shopping
7	Shopping	家電 新生活	Shopping
7	Shopping	家電 引っ越し	Shopping
7	Shopping	家電 電子レンジ	Shopping
7	Shopping	家電 一人暮らし	QA
8	Shopping	冷蔵庫	Shopping
8	Shopping	テレビ	News
8	Shopping	テレビ 値段	Shopping
8	Shopping	電子レンジ 値段 価格	Shopping
8	Shopping	ps4 slim	News
8	Shopping	扇風機	Shopping
9	Shopping	冷蔵庫	Shopping
9	Shopping	冷蔵庫 一人暮らし	Shopping
9	Shopping	家電 一人暮らし	Shopping

10	Shopping	冷蔵庫	Shopping
10	Shopping	冷蔵庫 一人	Shopping
10	Shopping	冷蔵庫 単身	Shopping
10	Shopping	冷蔵庫 単身 明後日	QA
10	Shopping	冷蔵庫 単身 速	QA
10	Shopping	冷蔵庫 単身 すぐ届く	QA
10	Shopping	家電 単身 すぐ届く	QA
10	Shopping	家電 あす楽	Shopping
10	Shopping	テレビ あす楽	Shopping
10	Shopping	電子レンジ あす楽	Shopping
10	Shopping	冷蔵庫 あす楽	Shopping
10	Shopping	洗濯機 あす楽	Shopping
1	QA	自己 PR	QA
1	QA	エントリーシート の 弱み	QA
1	QA	志望動機	QA
2	QA	エントリーシート 書き方	News
2	QA	エントリーシート 見本	QA
2	QA	エントリーシート 自己 PR	QA
2	QA	エントリーシート 強み	News
2	QA	エントリーシート 強み 書き方	News
2	QA	エントリーシート 強み 弱み	News
2	QA	エントリーシート 強み 弱み	News
2	QA	エントリーシート 強み 注意点	QA
3	QA	エントリーシート 書き方	News
3	QA	自己 pr	QA
3	QA	自己 pr 就活	QA
3	QA	エントリーシート の 弱み	QA
4	QA	エントリーシート	QA
4	QA	エントリーシート 良い例	QA
4	QA	エントリーシート 強み 弱み	News
4	QA	エントリーシート 注意点	QA
5	QA	エントリーシート 書き方	News
5	QA	エントリーシート 見本	QA
5	QA	エントリーシート 添削	QA
5	QA	エントリーシート 注意点	QA

6	QA	エントリーシート	QA
6	QA	エントリーシート 書き方	News
7	QA	エントリーシート	QA
7	QA	エントリーシート 書き方	News
7	QA	ES 書き方	QA
8	QA	エントリーシート	QA
8	QA	自分の弱み	QA
8	QA	就職	Shopping
8	QA	エントリーシート 書き方	News
9	QA	エントリーシート自己	QA
9	QA	エントリーシート自己 PR	QA
10	QA	就活 es	QA
10	QA	就活 es 注意点	News

対照群のクエリ：

参加者 ID	検索 課題	クエリ	ランダム選んだ 意図
11	News	熊本県庁	QA
11	News	熊本県 文化遺産	QA
11	News	熊本城	QA
11	News	熊本 橋 崩落	News
11	News	熊本 復興進まず	News
11	News	熊本 ボランティア活動	News
11	News	熊本 復興途上	News
12	News	熊本 地震 被害	News
12	News	熊本 復興 9 月	News
12	News	熊本 復興 9 月	Shopping
12	News	熊本 復興 進捗	News
12	News	熊本 地震 復興	News
12	News	熊本 地震 復興	Shopping
12	News	熊本地震 復興 現在	News
12	News	熊本地震 復興 現在	Shopping
12	News	熊本 復興 状況	Shopping
13	News	2016 4 14 地震	Shopping
13	News	2016 4 14 地震 写真	QA
13	News	2016 熊本地震 復興	News
13	News	2016 熊本地震 状況	QA
13	News	熊本	QA
13	News	熊本 地震	Shopping
13	News	熊本 地震 復興	News
13	News	熊本 地震 復興状況	News
14	News	熊本地震	News
14	News	熊本地震 被害	News
14	News	熊本地震 復興	Shopping
14	News	熊本地震 復興状況	QA
14	News	熊本地震 2016 被害	News
14	News	熊本地震 2016 被害 気象庁	QA
15	News	熊本地震 復興状況	QA
15	News	熊本 被害	News

16	News	熊本地震	Shopping
16	News	熊本地震 復興状況	QA
16	News	熊本地震 復興状況 9月	QA
16	News	熊本地震 復興 現状	News
16	News	熊本地震 復興 現状	QA
17	News	熊本 復興 5ヶ月	News
17	News	熊本地震 復興 5ヶ月	News
18	News	熊本地震 復興状況	QA
18	News	熊本 復興 インフラ	News
18	News	熊本地震 インフラ	Shopping
18	News	熊本地震 インフラ 11月	QA
18	News	熊本地震 11月	Shopping
18	News	熊本地震 復興作業	QA
18	News	熊本地震 復興作業	News
19	News	熊本地震 復興作業	Shopping
19	News	熊本地震 復興 報告	Shopping
19	News	熊本地震 5ヶ月	QA
19	News	熊本地震 5ヶ月	Shopping
20	News	熊本地震	Shopping
20	News	熊本地震 被害	News
11	Shopping	ケーズデンキ 筑波大学	News
11	Shopping	ケーズデンキ つくば研究学園店	News
11	Shopping	中古家電 つくば市	News
11	Shopping	一人暮らし 家電 茨城	Shopping
11	Shopping	家電 配送無料	Shopping
11	Shopping	家電 配送無料 茨城	News
11	Shopping	家電 即日配送 つくば	News
11	Shopping	冷蔵庫 小さい 目安	QA
11	Shopping	価格コム 電子レンジ	QA
11	Shopping	価格コム 電子レンジ	Shopping
11	Shopping	価格コム 薄型テレビ	Shopping
12	Shopping	一人暮らし 家電 相場	News
12	Shopping	冷蔵庫 格安	QA
12	Shopping	冷蔵庫 3万 2日後	Shopping
12	Shopping	冷蔵庫 店頭 家に届く	News

12	Shopping	冷蔵庫 店頭 配送	News
12	Shopping	冷蔵庫 つくば 配送	QA
12	Shopping	家電 配送 つくば	QA
12	Shopping	家電 つくば おすすめ	News
12	Shopping	一人暮らし	Shopping
12	Shopping	一人暮らし テレビ おすすめ	News
12	Shopping	一人暮らし テレビ おすすめ	QA
12	Shopping	家電 格安 比較	Shopping
13	Shopping	筑波大学 家電	News
13	Shopping	一人暮らし 大学生 家具	Shopping
13	Shopping	筑波大生 家具	News
13	Shopping	AMAZON	News
13	Shopping	価格コム	Shopping
13	Shopping	筑波大生 一人暮らし	QA
13	Shopping	筑波大生 一人暮らし アパート	News
13	Shopping	一人暮らし 大学生 家具	QA
13	Shopping	一人暮らし 大学生	News
13	Shopping	家具 サイズ	QA
13	Shopping	家具 サイズ 一人暮らし	News
13	Shopping	一人暮らし 家電 選び方	Shopping
14	Shopping	つくば 家電	Shopping
14	Shopping	つくば 冷蔵庫	News
14	Shopping	つくば 電気屋	News
14	Shopping	つくば ノジマ電器	QA
14	Shopping	冷蔵庫 一人暮らし	QA
14	Shopping	冷蔵庫 80L つくば	QA
14	Shopping	冷蔵庫 つくば	QA
15	Shopping	一人暮らし 家具	News
15	Shopping	一人暮らし 家具	QA
16	Shopping	家電 つくば	News
16	Shopping	ケーズデンキ 商品	Shopping
16	Shopping	ノジマ電器 つくば	News
17	Shopping	一人暮らし 家電	Shopping
17	Shopping	家電 配送 2日以内	QA
17	Shopping	家電セット 価格	News

17	Shopping	家電セット 価格 比較	Shopping
17	Shopping	家電セット 価格 比較	QA
17	Shopping	家電セット	News
17	Shopping	家電セット	QA
18	Shopping	つくば 家電 一人暮らし	News
18	Shopping	筑波大学 家電 一人暮らし	News
18	Shopping	筑波大学 家電購入 一人暮らし	QA
18	Shopping	筑波大学 リサイクル	News
18	Shopping	冷蔵庫 価格コム	Shopping
18	Shopping	冷蔵庫 価格コム	News
18	Shopping	冷蔵庫 一人暮らし オススメ	News
18	Shopping	冷蔵庫 一人暮らし オススメ	QA
18	Shopping	洗濯機 一人暮らし オススメ	News
18	Shopping	洗濯機 一人暮らし オススメ	Shopping
18	Shopping	洗濯機 一人暮らし ランキング	QA
18	Shopping	つくば 家電	News
18	Shopping	家電 一人暮らし	News
18	Shopping	家電 配達 費用	News
19	Shopping	冷蔵庫	QA
19	Shopping	家電 セット 明後日	News
19	Shopping	家電 セット 翌日配送	Shopping
19	Shopping	価格 com	News
19	Shopping	価格 com 家電セット	News
20	Shopping	家具	News
20	Shopping	amazon	QA
20	Shopping	yahoo	News
11	QA	ITpro	Shopping
11	QA	リクナビ	News
11	QA	エントリーシート	QA
11	QA	エントリーシート リクナビ	QA
11	QA	学情	QA
11	QA	人事が拾う履歴書	Shopping
11	QA	ダメな履歴書	Shopping
11	QA	Itpro 履歴書	News
11	QA	Buzzfeed 履歴書	News

11	QA	Buzzfeed 履歴書	Shopping
11	QA	Buzzfeed	News
11	QA	Ascii 履歴書	News
11	QA	Atmarkit エントリーシート	News
11	QA	Atmarkit エントリーシート	QA
11	QA	自己 PR 失敗例	Shopping
12	QA	就活 エントリーシート	QA
12	QA	就活 エントリーシート 書き方	News
12	QA	エントリーシート 書き方 強み	Shopping
12	QA	エントリーシート 書き方 弱み	News
12	QA	エントリーシート 書き方 自己 PR	QA
13	QA	エントリーシート 書き方	QA
13	QA	エントリーシート 自己 PR	News
13	QA	エントリーシート 自己 PR 具体例	QA
13	QA	エントリーシート 自己 PR 失敗	News
13	QA	エントリーシート 強み	News
14	QA	エントリーシート コツ	QA
14	QA	エントリーシート 読みやすい	Shopping
14	QA	エントリーシート 読みやすい	QA
14	QA	エントリーシート 注意点	News
14	QA	エントリーシート 自己 PR	Shopping
15	QA	エントリーシート 書き方	QA
15	QA	エントリーシート 書き方	News
16	QA	ES 書き方	QA
16	QA	エントリーシート 書き方	News
16	QA	エントリーシート 書き方 就活	News
16	QA	エントリーシート 書き方 就活	QA
17	QA	就活 エントリーシート 書き方	Shopping
17	QA	就活 エントリーシート 例	News
17	QA	就活 エントリーシート 注意点	News
17	QA	就活 エントリーシート 内容	QA
18	QA	エントリーシート 大手	QA
18	QA	エントリーシート 大手 例	News
18	QA	エントリーシート みんなのしゅうしょ	QA
18	QA	エントリーシート みんなの就職	News

18	QA	エントリーシート みんなの就職	QA
19	QA	自己 PR 書き方	News
19	QA	エントリーシート 例	QA
20	QA	PR	QA
20	QA	自己分析	QA
20	QA	sony	QA
20	QA	sony entry sheet	News